GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 ; Search time 1010.34 Seconds
(without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
Perfect score: 501
Sequence: 1 caatttgtggaagcttgcgg.....ttgtgtcttaacaaaggtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	489	97.6	835	9	AF188523	AF188523 AF188523
	2	483	96.4	695	9	AL133724	AL133724 DKFZp761J
	3	483	96.4	777	13	BI602169	BI602169 603246290
→ c	4	479.8	95.8	644	10	AW162535	AW162535 au77b01.x
	5	479	95.6	920	12	BG676871	BG676871 602623473
	6	437.2	87.3	801	13	BI603180	BI603180 603249824
c	7	427.2	85.3	546	12	BF748808	BF748808 MR0-BN011
	8	410.8	82.0	679	10	AW953240	AW953240 EST365310
	9	385.2	76.9	574	12	BF667219	BF667219 602121018
	10	369	73.7	521	13	BM353207	BM353207 ig44a10.y
	11	368	73.5	539	13	BM353362	BM353362 ig46a10.y
→	12	359	71.7	359	9	AA484945	AA484945 aa41b06.r
c	13	311.4	62.2	609	13	BG992490	BG992490 MR3-HT103
	14	301	60.1	870	12	BG425739	BG425739 602492044
	15	299	59.7	354	9	AI372818	AI372818 EST175362
	16	293.8	58.6	657	10	BB309170	BB309170 BB309170
c	17	282.8	56.4	359	9	AA705585	AA705585 ab42b01.s
	18	272.6	54.4	371	9	AA151055	AA151055 zl45c03.r
	19	271	54.1	310	9	AI372817	AI372817 EST175361
	20	269.8	53.9	483	12	BG384511	BG384511 303736 MA
	21	263.8	52.7	523	10	BE372746	BE372746 601224258
	22	254	50.7	896	12	BG468527	BG468527 602510307
	23	251	50.1	501	12	BF044659	BF044659 BP250020A
	24	250	49.9	658	13	BI153873	BI153873 602870875
	25	239.6	47.8	374	10	AW484613	AW484613 61734 MAR
c	26	237.8	47.5	371	9	AU083157	AU083157 AU083157
	27	233	46.5	388	13	BI050739	BI050739 RC6-GN007
	28	225.6	45.0	377	12	BF042085	BF042085 BP250013A
	29	217	43.3	491	14	BQ694790	BQ694790 1000933 H
	30	216.4	43.2	594	10	AV720937	AV720937 AV720937
	31	207.6	41.4	765	9	AL553184	AL553184 AL553184
	32	207.2	41.4	328	9	AA879884	AA879884 vw01c06.r
	33	206.8	41.3	394	13	BI337284	BI337284 AR088C04L
	34	205	40.9	749	12	BF695548	BF695548 601851957
	35	203.8	40.7	450	12	BE703924	BE703924 MR0-NN112
c	36	202.2	40.4	352	14	T03019	T03019 FB20D6 Feta
	37	192.2	38.4	334	12	BF041935	BF041935 BP250006A
c	38	188	37.5	1039	9	AL570666	AL570666 AL570666
	39	158.4	31.6	333	10	AW463757	AW463757 BP230013B
	40	156.2	31.2	370	12	BE930535	BE930535 RC6-GN007
	41	149.2	29.8	321	10	AW227478	AW227478 up10h03.y
	42	146.4	29.2	535	12	BF963798	BF963798 QV2-NN004
	43	145.4	29.0	184	12	BE699344	BE699344 RC3-NN006
c	44	143	28.5	512	12	BF090072	BF090072 QV2-NN004
	45	141.6	28.3	705	13	BG922853	BG922853 602817405

ALIGNMENTS

RESULT 1
AF188523/c
LOCUS AF188523 835 bp mRNA linear EST 02-MAR-2000
DEFINITION AF188523 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens cDNA clone ISG 4, mRNA sequence.
ACCESSION AF188523
VERSION AF188523.1 GI:7144569
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS Ye,Z. and Connor,J.R.
TITLE Identification of Iron Regulated Genes by Rescreening cDNA Libraries from SSH with Antisense Probe from Three Iron Conditions
JOURNAL Unpublished (2000)
COMMENT Contact: Ye Z
Neuroscience and Anatomy
Pennsylvania State University College of Medicine
500 University Drive, Hershey, PA 17033, USA
library screened by SSH and reverse Northern blot; increased expression in iron loading was confirmed by Northern blot.
FEATURES Location/Qualifiers
source 1. .835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ISG 4"
/clone_lib="Homo sapiens ATCC HTB-12; SW1088"

```
                     /tissue_type="astrocytoma"
                     /cell_line="ATCC HTB-12; SW1088"
                     /note="Organ: liver"
BASE COUNT      225 a    209 c    156 g    245 t
ORIGIN

  Query Match             97.6%;  Score 489;  DB 9;  Length 835;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-117;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 710

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 650

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 590

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 530

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     529 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 470

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 410

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 350

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     349 AGGAGTTGTAAGAAGGTGCACATA-TTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 291

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db     290 ATTGTGTCTTAACAAAGGTGT 270


RESULT 2
AL133724
LOCUS       AL133724                695 bp    mRNA    linear   EST 25-FEB-2000
DEFINITION  DKFZp761J19121_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
            DKFZp761J19121 5', mRNA sequence.
ACCESSION   AL133724
VERSION     AL133724.1  GI:6601912
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 695)
  AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Poustka, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Poustka A.J.
            Department Lehrach
            Max-Planck-Institute for Molecular Genetics
            Ihnestrasse 73, 14195 Berlin, Germany
            Tel: +49-30-84131623
            Fax: +49-30-84131128
            Email: poustka@mpimg-berlin-dahlem.mpg.de
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp761J19121) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .695
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp761J19121"
                      /clone_lib="761 (synonym: hamy2)"
                      /tissue_type="amygdala"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      202 a    134 c    185 g    174 t
ORIGIN

  Query Match             96.4%;  Score 483;  DB 9;  Length 695;
  Best Local Similarity   98.4%;  Pred. No. 1.2e-115;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 72

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 132

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
           |||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db     133 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 192

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     193 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 252

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 312

Qy     293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     313 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 372

Qy     353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 432

Qy     413 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     433 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 492

Qy     473 AAGTTTTCATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||||||||||
Db     493 AAGTTTTCATTGTGTCTTAACAAAGGTGT 521


RESULT 3
BI602169
LOCUS       BI602169                777 bp    mRNA    linear   EST 07-SEP-2001
DEFINITION  603246290F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5288423 5',
            mRNA sequence.
ACCESSION   BI602169
VERSION     BI602169.1  GI:15495108
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 777)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11729  row: d  column: 24
            High quality sequence stop: 777.
FEATURES             Location/Qualifiers
     source          1. .777
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5288423"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      223 a    159 c    207 g    188 t
ORIGIN
```

```
  Query Match             96.4%;  Score 483;  DB 13;  Length 777;
  Best Local Similarity   98.4%;  Pred. No. 1.2e-115;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 267

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 327

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
           |||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db     328 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 387

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 447

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 507

Qy     293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 567

Qy     353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     568 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 627

Qy     413 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 687

Qy     473 AAGTTTTCATTGTGTCTTAACAAAGGTGT 501
```

```
           ||||||||||||||||||||||||||||
Db     688 AAGTTTTCATTGTGTCTTAACAAAGGTGT 716
```

```
RESULT 4
AW162535/c
LOCUS       AW162535                644 bp    mRNA    linear   EST 09-NOV-1999
DEFINITION  au77b01.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
            IMAGE:2782249 3' similar to TR:Q63603 Q63603 TRG MRNA ;, mRNA
            sequence.
ACCESSION   AW162535
VERSION     AW162535.1  GI:6301568
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: au77b01.y1
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2782249"
                     /clone_lib="Schneider fetal brain 00004"
                     /sex="male"
                     /tissue_type="frontal lobe"
                     /dev_stage="5 months post-conception"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescript SK (Stratagene);
                     Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
                     prepared from human fetal brain tissue.  5' and 3'
                     adaptors were used in cloning as follows: 5' adaptor
                     sequence:
                     5'-GAGAGAGAAGAGCTCAAGGATCCTTAATTAAATTAATCCCCCCCCCCCCC-3'
                     and 3' adaptor sequence:
                     5'-GAGAGAGACTCGAGTTTTTTTTTTTTTTTT-3'.  The library was
                     size-selected for >0.5 kb inserts and has an average
                     insert size estimated at 1.2 kb.  This library was
                     constructed using the CAP-trapper method for full-length
                     enrichment and has not undergone amplification.    Library
                     was constructed by Dr. Claudio Schneider (LNCIB-Area
                     Science Park,  Trieste, Italy). "
BASE COUNT      162 a    168 c    128 g    186 t
ORIGIN
```

```
  Query Match             95.8%;  Score 479.8;  DB 10;  Length 644;
  Best Local Similarity   98.0%;  Pred. No. 8.3e-115;
  Matches  499;  Conservative    0;  Mismatches    2;  Indels    8;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     638 CAATTTGTGCAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 579

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db    578 CAGCTCGAGTATCAGGAAGAAATGATAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 519

Qy    121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
          |||||||||||||||||        ||||||||||||||||||||||||||||||||||||
Db    518 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 459

Qy    173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    458 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 399

Qy    233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    398 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 339

Qy    293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 279

Qy    353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    278 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 219

Qy    413 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    218 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 159

Qy    473 AAGTTTTCATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||||||||||
Db    158 AAGTTTTCATTGTGTCTTAACAAAGGTGT 130
```

RESULT 5
BG676871
LOCUS BG676871 920 bp mRNA linear EST 01-MAY-2001
DEFINITION 602623473F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748442 5', mRNA sequence.
ACCESSION BG676871
VERSION BG676871.1 GI:13908268
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10600 row: a column: 19
High quality sequence stop: 849.

FEATURES Location/Qualifiers
source 1. .920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4748442"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 274 a 164 c 230 g 251 t 1 others
ORIGIN

Query Match 95.6%; Score 479; DB 12; Length 920;
Best Local Similarity 98.6%; Pred. No. 1.4e-114;
Matches 494; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
Qy      1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          |||||||||||||||    |   ||||||||||||||||||||||||||||||||||||
Db     19 CAATTTGTGGAAGCT--GCGGTCAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 76

Qy     61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 136

Qy    121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 196

Qy    181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 256

Qy    241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 316

Qy    301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 376

Qy    361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 436

Qy    421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 496

Qy    481 ATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||
Db    497 ATTGTGTCTTAACAAAGGTGT 517
```

RESULT 6
BI603180
LOCUS BI603180 801 bp mRNA linear EST 07-SEP-2001
DEFINITION 603249824F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301749 5', mRNA sequence.
ACCESSION BI603180
VERSION BI603180.1 GI:15496119
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11763 row: p column: 06
High quality sequence stop: 710.

FEATURES Location/Qualifiers
source 1. .801
/organism="Homo sapiens"
/db_xref="taxon:9606"

```
                     /clone="IMAGE:5301749"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', 
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      233 a    164 c    215 g    189 t
ORIGIN

  Query Match             87.3%;  Score 437.2;  DB 13;  Length 801;
  Best Local Similarity   96.3%;  Pred. No. 1.1e-103;
  Matches  493;  Conservative    0;  Mismatches    8;  Indels   11;  Gaps    4;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 267

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 327

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
           |||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db     328 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 387

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 447

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGAC-TT 291
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     448 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTT 507

Qy     292 GCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCAC 351
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 GCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCAC 567

Qy     352 AGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAG 411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     568 AGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAG 627

Qy     412 ACTTTCTATAGGAGTTGTAAGAAGGTGCA-CATATTTTTTTAAATCTCA-CTGGCAATAT 469
           |||||||||||||||||||||||||||||| |||||||||||||||| ||||||||||
Db     628 ACTTTCTATAGGAGTTGTAAGAAGGTGCACCATATTTTTTTAAATCTCACCTGGCAATAT 687

Qy     470 TCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||      | | |||||||||
Db     688 TCAAAGTTTTCATGTGTCTTAACCAAAGGTGT 719


RESULT 7
BF748808/c
LOCUS       BF748808                 546 bp    mRNA    linear   EST 10-JAN-2001
DEFINITION  MR0-BN0115-041000-013-e10 BN0115 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF748808
VERSION     BF748808.1  GI:12075471
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR0&t2=MR0-BN0115-
            041000-013-e10&t3=2000-10-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 16
            High quality sequence stop: 545.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BN0115"
                     /dev_stage="Adult"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      140 a    135 c    115 g    156 t
ORIGIN

  Query Match             85.3%;  Score 427.2;  DB 12;  Length 546;
  Best Local Similarity   99.1%;  Pred. No. 4.5e-101;
  Matches  440;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      59 ACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTT 118
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 ACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTT 487

Qy     119 CTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATT 178
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 CTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATT 427

Qy     179 CCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGA 238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 CCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGA 367

Qy     239 CCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGT 298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 CCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGT 307

Qy     299 CATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGG 358
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     306 CATTTGCAAACTCAGGATGCTTTCCAAAGCCCATCACTGGGGAGACCGAGCACAGGGAGG 247

Qy     359 ACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTT-CTTAACAGACTTTC 417
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     246 ACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCCTTAACAGACTTTC 187

Qy     418 TATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTT 477
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     186 TATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCCATATTCAAAGTT 127

Qy     478 TTCATTGTGTCTTAACAAAGGTGT 501
           ||||||||||||||| |||||||
Db     126 TTCATTGTGTCTTAACCAAGGTGT 103
```

```
RESULT 8
AW953240
LOCUS       AW953240                679 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST365310 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW953240
VERSION     AW953240.1  GI:8142923
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 48
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGB"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      194 a    146 c    191 g    147 t      1 others
ORIGIN

  Query Match           82.0%;  Score 410.8;  DB 10;  Length 679;
  Best Local Similarity  97.5%;  Pred. No. 8.8e-97;
  Matches  430;  Conservative    0;  Mismatches    3;  Indels    8;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     239 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 298

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 358

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
             |||||||||||||||||        ||||||||||||||||||||||||||||||||||||
Db     359 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 418

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 478

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 538

Qy     293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 598

Qy     353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
             |||||||||||||||||||||||||||| |||||||||||||||||||| ||||||||
Db     599 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGNAATAAAGAACAACGTTATTTTTTAACAGA 658

Qy     413 CTTTCTATAGGAGTTGTAAGA 433
             |||||| ||||||||||||||
Db     659 CTTTCTTTAGGAGTTGTAAGA 679
```

```
RESULT 9
BF667219
LOCUS       BF667219                574 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  602121018F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278077 5',
            mRNA sequence.
ACCESSION   BF667219
VERSION     BF667219.1  GI:11941114
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1101  row: c  column: 06
            High quality sequence stop: 544.
FEATURES             Location/Qualifiers
     source          1. .574
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4278077"
                     /clone_lib="NIH_MGC_56"
                     /tissue_type="primitive neuroectoderm"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.65 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      165 a    133 c    158 g    118 t
ORIGIN

  Query Match           76.9%;  Score 385.2;  DB 12;  Length 574;
  Best Local Similarity  99.2%;  Pred. No. 4.2e-90;
  Matches  387;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 ATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCA 62
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 AATTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCA 244

Qy      63 GCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGA 122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 GCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGA 304

Qy     123 AATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCT 182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 AATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCT 364

Qy     183 TCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAG 242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 TCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAG 424

Qy     243 CTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATT 302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 CTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATT 484
```

```
Qy     303 TGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCA 362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 TGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCA 544

Qy     363 AGGGGAAGGGGAGAGAAAGGAAATAAAGAA 392
           ||||||||||||||| |||||||| ||||
Db     545 AGGGGAAGGGGAGAGACAGGAAATACAGAA 574


RESULT 10
BM353207
LOCUS       BM353207                521 bp    mRNA    linear   EST 07-JAN-2002
DEFINITION  ig44a10.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q63603
            Q63603 TRG PROTEIN ;, mRNA sequence.
ACCESSION   BM353207
VERSION     BM353207.1  GI:18085565
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig44a10.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 462.
FEATURES             Location/Qualifiers
     source          1. .521
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      151 a    120 c    137 g    113 t
ORIGIN

  Query Match             73.7%;  Score 369;  DB 13;  Length 521;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-86;
  Matches  369;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 212

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 272

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 332

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 392

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 452

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 512

Qy     361 CAAGGGGAA 369
           |||||||||
Db     513 CAAGGGGAA 521


RESULT 11
BM353362
LOCUS       BM353362                539 bp    mRNA    linear   EST 07-JAN-2002
DEFINITION  ig46a10.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q63603
            Q63603 TRG PROTEIN ;, mRNA sequence.
ACCESSION   BM353362
VERSION     BM353362.1  GI:18085720
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 478.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
```

```
                    Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                    314-362-1916, Fax: 314-747-2692."
BASE COUNT      153 a    127 c    144 g    114 t      1 others
ORIGIN

  Query Match             73.5%;  Score 368;  DB 13;  Length 539;
  Best Local Similarity   99.7%;  Pred. No. 1.3e-85;
  Matches  368;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 230

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 290

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 350

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 410

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 470

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 530

Qy     361 CAAGGGGAA 369
           ||||| |||
Db     531 CAAGGNGAA 539


RESULT 12
AA484945
LOCUS       AA484945                359 bp    mRNA    linear   EST 15-AUG-1997
DEFINITION  aa41b06.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:815795 5'
            similar to TR:G550420 G550420 TRG MRNA ;, mRNA sequence.
ACCESSION   AA484945
VERSION     AA484945.1  GI:2214164
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 359)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .359
                     /organism="Homo sapiens"
                     /db_xref="GDB:6035413"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:815795"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      108 a     80 c     97 g     74 t
ORIGIN

  Query Match             71.7%;  Score 359;  DB 9;  Length 359;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-83;
  Matches  359;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      47 TGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGG 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGG 60

Qy     107 CGAAGGAGCTTTCTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCG 166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGAAGGAGCTTTCTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCG 120

Qy     167 TCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGG 226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGG 180

Qy     227 TTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGG 286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGG 240

Qy     287 GACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCG 346
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCG 300

Qy     347 AGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCT 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCT 359


RESULT 13
BG992490/c
LOCUS       BG992490                609 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  MR3-HT1039-020201-002-b09 HT1039 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG992490
VERSION     BG992490.1  GI:14396560
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 609)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
```

```
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-HT1039-
            020201-002-b09&t3=2001-02-02&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 6
            High quality sequence stop: 605.
FEATURES              Location/Qualifiers
     source           1. .609
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="HT1039"
                      /dev_stage="Adult"
                      /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
BASE COUNT      133 a    160 c    157 g    158 t      1 others
ORIGIN

  Query Match             62.2%;  Score 311.4;  DB 13;  Length 609;
  Best Local Similarity   97.6%;  Pred. No. 8e-71;
  Matches  326;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 CCATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 275

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     274 CAGCTCGAGTATCAGGAAGAAATGAAAGCCCACTACAGGGAAATGGCGAAGGAGCTTTCT 215

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 155

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     154 CTTCACATCTTCAACGCCATCAGTGGGACTCCNACAAGCACAATGGTTCACGGGATGACC 95

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      94 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGCGTGGGGACTTGCTTTGTCA 35

Qy     301 T-TTGCAAACTCAGGATGCTTTCCAAAGCCAATC 333
           | ||||||  |||||||||||||| |||||||||
Db      34 TGTTGCAACGTCAGGATGCTTTCCCAAGCCAATC 1


RESULT 14
BG425739
LOCUS       BG425739                 870 bp    mRNA    linear   EST 14-MAR-2001
DEFINITION  602492044F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4606114 5',
            mRNA sequence.
ACCESSION   BG425739
VERSION     BG425739.1  GI:13332245
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1344  row: g  column: 11
            High quality sequence stop: 617.
FEATURES              Location/Qualifiers
     source           1. .870
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4606114"
                      /clone_lib="NIH_MGC_75"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                      SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                      3' adaptors were used in cloning as follows: 5' adaptor
                      sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                      5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                      C, or G and N = A, C,  G, or T).  Average insert size 1.65
                      kb (range 0.5-4.0 kb). 15/15 colonies contained  inserts
                      by PCR. This library was enriched for full-length clones
                      and was constructed by Clontech  Laboratories (Palo Alto,
                      CA). Note: this is a NIH_MGC Library."
BASE COUNT      231 a    205 c    197 g    237 t
ORIGIN

  Query Match             60.1%;  Score 301;  DB 12;  Length 870;
  Best Local Similarity   95.4%;  Pred. No. 4.3e-68;
  Matches  354;  Conservative    0;  Mismatches   10;  Indels    7;  Gaps    4;

Qy     137 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 215

Qy     197 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     216 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 275

Qy     257 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 335

Qy     317 GCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAA--GGGGA 374
           |||||||||||||||||||| |||||||||||||||||||| ||||||  ||| |
Db     336 GCTTTCCAAAGCCAATCACTGGGCAGACCGAGCACAGGGAGGACCACGGGGAACGGGGAG 395

Qy     375 GAGAAAGGAAATAA-AGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGA 433
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     396 CAGAAAGGAAATAACAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGA 455

Qy     434 AGGTGCACATATTT---TTTTAAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTT 490
           |||||||||||||    || | |||||||||||||||||| ||||| || ||||||||
Db     456 AGGTGCACATATTTCTCCTTAACATCTCACTGGCAATATTCACAGTTTCCA-TGTGTCTT 514

Qy     491 AACAAAGGTGT 501
           ||||||||||| 
Db     515 AACAAAGGTGT 525


RESULT 15
AI372818
LOCUS       AI372818                 354 bp    mRNA    linear   EST 12-JAN-1999
DEFINITION  EST175362 Infant brain, Bento Soares Homo sapiens cDNA clone
            FUHIBU1, mRNA sequence.
ACCESSION   AI372818
VERSION     AI372818.1  GI:4152684
KEYWORDS    EST.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
            EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
            EST175339 EST17534
            Contact: HGI (Human Gene Index)
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-0200
            Fax: (301)-838-0208
            Email: hgi@tigr.org.
FEATURES              Location/Qualifiers
     source           1. .354
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="FUHIBU1"
                      /clone_lib="Infant brain, Bento Soares"
                      /lab_host="E. coli DH5-alpha"
                      /note="Vector: BA, M13-derived; Site_1: HindIII; Site_2:
                      NotI; The infant brain library, constructed by Bento
                      Soares, Columbia University, was oligo-(dT) primed and
                      directionally cloned into an M13-derived plasmid using
                      total brain mRNA from a 72-day old human female afflicted
                      with spinal muscular atrophy. "
BASE COUNT       97 a     77 c    100 g     79 t      1 others
ORIGIN

  Query Match             59.7%;  Score 299;  DB 9;  Length 354;
  Best Local Similarity   94.4%;  Pred. No. 1.3e-67;
  Matches  334;  Conservative    0;  Mismatches   11;  Indels    9;  Gaps    2;

Qy    100 GAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCA--------GATCTGCCCCCTGG 151
          ||||||||||||||||||||||||||||||||||||        ||||||||||||||
Db      1 GAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGG 60

Qy    152 AGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTC 211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTC 120

Qy    212 CAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCA 271
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCA 180

Qy    272 TGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAA 331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAA 240

Qy    332 TCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGA 391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGA 300

Qy    392 ACAACGTTATTTCTTAACA-GACTTTCTATAGGAGTTGTAAGAAGGTGCACATA 444
          ||||||||||||||||||| ||||||| |||||||||     | | |  || |
Db    301 ACAACGTTATTTCTTAACAGGACTTTCTTTAGGAGTTGGTNAGAAGGGTGCACA 354

Search completed: February  7, 2003, 08:49:46
Job time : 1017.34 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 994.674 Seconds
(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttttacaccatcaccaaa.....aggccacgtgatgatcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		498.4	99.7	7506	6	AX255048	AX255048 Sequence
2		498.4	99.7	7522	9	AF527605	AF527605 Homo sapi
3		498.4	99.7	7545	9	AB028981	AB028981 Homo sapi
4		204.4	40.9	6454	6	AX173022	AX173022 Sequence
5		194.8	39.0	139887	9	CNS01RGX	AL160233 Human chr
6	c	126.8	25.4	163316	9	AL161420	AL161420 Human DNA
7	c	90.4	18.1	175281	2	AC109966	AC109966 Rattus no
8		85	17.0	4391	6	AX173028	AX173028 Sequence
9		85	17.0	4393	6	AX173118	AX173118 Sequence
10		75.8	15.2	175281	2	AC109966	AC109966 Rattus no
11	c	75.8	15.2	192825	2	AC119357	AC119357 Rattus no
12		71.4	14.3	5564	9	AK074081	AK074081 Homo sapi
13		71.4	14.3	6422	9	AK090429	AK090429 Homo sapi
14		71.4	14.3	7215	6	AX174569	AX174569 Sequence
15		71.4	14.3	9298	9	HSM803577	AL832270 Homo sapi
16	c	69.8	14.0	665	6	AX371133	AX371133 Sequence
17		60.8	12.2	181127	10	AL672038	AL672038 Mouse DNA
18		60.2	12.0	58723	9	AL391237	AL391237 Human DNA
19	c	59	11.8	162378	2	AC105541	AC105541 Rattus no
20	c	59	11.8	174007	2	AC118998	AC118998 Rattus no
21		58.2	11.6	82603	2	AC111850	AC111850 Rattus no
22	c	52.8	10.6	171811	9	AC011739	AC011739 Homo sapi
23		52	10.4	181179	9	AL161725	AL161725 Human DNA
24	c	48.4	9.7	167215	5	AY016023	AY016023 Sphoeroid
25	c	48	9.6	281	11	G47605	G47605 Z25470_1 Ze
26		46	9.2	4886	9	AB037816	AB037816 Homo sapi
27		46	9.2	6372	6	AX173175	AX173175 Sequence
28	c	43.4	8.7	82603	2	AC111850	AC111850 Rattus no
29		42.6	8.5	147556	2	AC110359	AC110359 Rattus no
30	c	42.4	8.5	73515	2	AC017375	AC017375 Drosophil
31		42.4	8.5	166626	3	AC008318	AC008318 Drosophil
32		42.4	8.5	261475	2	AC096357	AC096357 Rattus no
33		42.4	8.5	301639	3	AE003590	AE003590 Drosophil
34		41	8.2	6828	6	AX172874	AX172874 Sequence
35		40.4	8.1	211	6	AX173251	AX173251 Sequence
36		40.4	8.1	3023	9	AK022412	AK022412 Homo sapi
37	c	40.4	8.1	122321	9	AC011472	AC011472 Homo sapi
38	c	38.6	7.7	69694	2	AC020073	AC020073 Drosophil
39	c	38.6	7.7	156222	9	AC087224	AC087224 Homo sapi
40		38.6	7.7	159108	2	AC026031	AC026031 Homo sapi
41	c	38.6	7.7	159350	3	AC008354	AC008354 Drosophil
42	c	38.6	7.7	257224	3	AE003617	AE003617 Drosophil
43	c	37.8	7.6	181994	2	AC116703	AC116703 Mus muscu
44	c	37.6	7.5	154846	2	AP001374	AP001374 Homo sapi
45		37.6	7.5	156211	9	AC091551	AC091551 Homo sapi

ALIGNMENTS

RESULT 1
AX255048
LOCUS AX255048 7506 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 7 from Patent WO0170808.
ACCESSION AX255048
VERSION AX255048.1 GI:16074541
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7506)
AUTHORS Rastelli,L.K. and Gerritsen,M.
TITLE Angiogenesis-associated proteins, and nucleic acids encoding the same

```
  JOURNAL   Patent: WO 0170808-A 7 27-SEP-2001;
            Curagen Corporation (US) ; GENENTECH, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .7506
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2147 a   1689 c   1764 g   1906 t
ORIGIN

  Query Match           99.7%;  Score 498.4;  DB 6;  Length 7506;
  Best Local Similarity 99.8%;  Pred. No. 3.7e-146;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2142 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 2201

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2202 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 2261

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2262 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 2321

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2322 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 2381

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2382 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 2441

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2442 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 2501

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2502 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 2561

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2562 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 2621

Qy     481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    2622 AGGCCACGTGATGATCGCCT 2641


RESULT 2
AF527605
LOCUS       AF527605                7522 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens zizimin1 mRNA, complete cds.
ACCESSION   AF527605
VERSION     AF527605.1  GI:22038158
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7522)
  AUTHORS   Meller,N., Irani-Tehrani,M., Kiosses,W.B., Del Pozo,M.A. and
            Schwartz,M.A.
  TITLE     Zizimin1, a novel Cdc42 activator, reveals new guanine nucleotide
            exchange-exchange factor domain for rho proteins
  JOURNAL   Nat. Cell Biol. (2002) In press
REFERENCE   2  (bases 1 to 7522)
  AUTHORS   Meller,N. and Schwartz,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2002) Cell Biology, The Scripps Research
            Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .7522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
     CDS             56. .6265
                     /note="guanine nucleotide exchange factor; Cdc42
                     activator"
                     /codon_start=1
                     /product="zizimin1"
                     /protein_id="AAM90306.1"
                     /db_xref="GI:22038159"
                     /translation="MSQPPLLPASAETRKFTRALSKPGTAAELRQSVSEVVRGSVLLA
                     KPKLIEPLDYENVIVQKKTQILNDCLREMLLFPYDDFQTAILRRQGRYICSTVPAKAE
                     EEAQSLFVTECIKTYNSDWHLVNYKYEDYSGEFRQLPNKVVKLDKLPVHVYEVDEEVD
                     KDEDAASLGSQKGGITKHGWLYKGNMNSAISVTMRSFKRRFFHLIQLGDGSYNLNFYK
                     DEKISKEPKGSIFLDSCMGVVQNNKVRRFAFELKMQDKSSYLLAADSEVEMEEWITIL
                     NKILQLNFEAAMQEKRNGDSHEDDEQSKLEGSGSGLDSYLPELAKSAREAEIKLKSES
                     RVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGKRILVKCNDLSFNLQCCVAENEEGPTT
                     NVEPFFVTLSLFDIKYNRKISADFHVDLNHFSVRQMLATTSPALMNGSGQSPSVLKGI
                     LHEAAMQYPKQGIFSVTCPHPDIFLVARIEKVLQGSITHCAEPYMKSSDSSKVAQKVL
                     KNAKQACQRLGQYRMPFAWAARTLFKDASGNLDKNARFSAIYRQDSNKLSNDDMLKLL
                     ADFRKPEKMAKLPVILGNLDITIDNVSSDFPNYVNSSYIPTKQFETCSKTPITFEVEE
                     FVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQKSFAKARNIAICIEFKDSDEEDSQPLK
                     CIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEIKIELPTQLHEKHHLLLTFFHVSCDNS
                     SKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHIPVSANLPSGYLGYQELGMGRHYG
                     PEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYCQKTESGAQALGNELVKYLKSL
                     HAMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVIIHVVAQCHEEGLESHLRSY
                     VKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNKLLKYSWFFFDVLIKSMA
                     QHLIENSKVKLLRNQRFPASYHHAVETVVNMLMPHITQKFRDNPEASKNANHSLAVFI
                     KRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNHEHYIPLNLPMPFG
                     KGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLREVGTALQEFREVRLIAISVLKNLLIK
                     HSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMTVKDESLALP
                     AVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGSLISTDSG
                     NSLPERNSEKSNSLDKHQQSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALF
                     TYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARTGMMHARLQQLGSLDNSLTFNHS
                     YGHSDADVLHQSLLEANIATEVCLTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVY
                     LCFLQKHQSETALKNVFTALRSLIYKFPSTFYEGRADMCAALCYEILKCCNSKLSSIR
                     TEASQLLYFLMRNNFDYTGKKSFVRTHLQVIISVSQLIADVVGIGGTRFQQSLSIINN
                     CANSDRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHENDPEMLVDLQYSLAKSYAST
                     PELRKTWLDSMARIHVKNGDLSEAAMCYVHVTALVAEYLTRKEAVQWEPPLLPHSHSA
                     CLRRSRGGVFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVLMELLEQCADGLW
                     KAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVMHSGRRLLGTYFR
                     VAFFGQAAQYQFTDSETDVEGFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGS
                     ENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEM
                     PFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKK
                     VAELRQLCSSAEVDMIKLQLKLQGSVSVQVNAGPLAYARAFLDDTNTKRYPDNKVKLL
                     KEVFRQFVEACGQALAVNERLIKEDQLEYQEEMKANYREMAKELSEIMHEQLG"
BASE COUNT     2146 a   1694 c   1774 g   1908 t
ORIGIN

  Query Match           99.7%;  Score 498.4;  DB 9;  Length 7522;
  Best Local Similarity 99.8%;  Pred. No. 3.7e-146;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2155 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 2214

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2215 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 2274

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2275 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 2334

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2335 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 2394

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2395 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 2454
```

```
Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2455 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 2514

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2515 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 2574

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2575 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 2634

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db    2635 AGGCCACGTGATGATCGCCT 2654
```

```
RESULT 3
AB028981
LOCUS       AB028981               7545 bp    mRNA    linear   PRI 10-MAY-2002
DEFINITION  Homo sapiens mRNA for KIAA1058 protein, partial cds.
ACCESSION   AB028981
VERSION     AB028981.2  GI:20521745
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh12146s1.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
  MEDLINE   99397452
   PUBMED   10470851
REFERENCE   2  (bases 1 to 7545)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
COMMENT     On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES             Location/Qualifiers
     source          1. .7545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="hh12146s1"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
                     /note="This sequence was obtained bysubcloning of the DNA
                     fragments derived from two cDNA clones (1 - 1846 was
                     derived from ef00877 and 1847 - 7545 was derived from
                     hh12146)."
     gene            1. .7545
                     /gene="KIAA1058"
     CDS             <1. .6288
                     /gene="KIAA1058"
                     /codon_start=1
                     /product="KIAA1058 protein"
                     /protein_id="BAA83010.2"
                     /db_xref="GI:20521746"
                     /translation="LACLCLARPGTRSCSSCRLSDLRFQPRMQADKCRTSSRSVKKEL
                     VIESPLQYKDAAQGEVEAESPGPVPAKPKLIEPLDYENVIVQKKTQILNDCLREMLLF
                     PYDDFQTAILRRQGRYICSTVPAKAEEEAQSLFVTECIKTYNSDWHLVNYKYEDYSGE
                     FRQLPNKVVKLDKLPVHVYEVDEEVDKDEDAASLGSQKGGITKHGWLYKGNMNSAISV
                     TMRSFKRRFFHLIQLGDGSYNLNFYKDEKISKEPKGSIFLDSCMGVVQNNKVRRFAFE
                     LKMQDKSSYLLAADSEVEMEEWITILNKILQLNFEAAMQEKRNGDSHEDDEQSKLEGS
                     GSGLDSYLPELAKSAREAEIKLKSESRVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGK
                     RILVKCNDLSFNLQCCVAENEEGPTTNVEPFFVTLSLFDIKYNRKISADFHVDLNHFS
                     VRQMLATTSPALMNGSGQSPSVLKGILHEAAMQYPKQGIFSVTCPHPDIFLVARIEKV
                     LQGSITHCAEPYMKSSDSSKVAQKVLKNAKQACQRLGQYRMPFAWAARTLFKDASGNL
                     DKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSDFPN
                     YVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQK
                     SFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEI
                     KIELPTQLHEKHHLLLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSE
                     QHIPVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHN
                     FFQYCQKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLTRATQEEVA
                     VNVTRVIIHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKP
                     SADFLTSNKLLKYSWFFFDVLIKSMAQHLIENSKVKLLRNQRFPASYHHAVETVVNML
                     MPHITQKFRDNPEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLF
                     EYKFEFLRVVCNHEHYIPLNLPMPFGKGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLR
                     EVGTALQEFREVRLIAISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQ
                     RINVRDVSPFPVNAGMTVKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASP
                     YTTSTPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSSTLGNSVVRCDK
                     LDQSEIKSLLMCFLYILKSMSDDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRY
                     IARTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSL
                     FTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRSLIYKFPSTFY
                     EGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSFVRTHLQVII
                     SVSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKRIRTVLMATA
                     QMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSEAAMCYVHVT
                     ALVAEYLTRKEAVQWEPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPNIDEEASMME
                     DVGMQDVHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLY
                     DTLHRAYSKVTEVMHSGRRLLGTYFRVAFFGQAAQYQFTDSETDVEGFFEDEDGKEYI
                     YKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPF
                     FDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVK
                     KRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLCSSAEVDMIKLQLKLQGSVSVQVNA
                     GPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGQALAVNERLIKEDQLEYQEE
                     MKANYREMAKELSEIMHEQLG"
BASE COUNT     2175 a   1687 c   1753 g   1930 t
ORIGIN
```

```
  Query Match           99.7%;  Score 498.4;  DB 9;  Length 7545;
  Best Local Similarity  99.8%;  Pred. No. 3.7e-146;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2178 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 2237

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2238 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 2297

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2298 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 2357

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2358 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 2417

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    2418 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 2477

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2478 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 2537

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2538 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 2597

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2598 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 2657

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db    2658 AGGCCACGTGATGATCGCCT 2677
```

```
RESULT 4
AX173022
LOCUS       AX173022               6454 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 1 from Patent WO0142294.
ACCESSION   AX173022
VERSION     AX173022.1  GI:14597975
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6454)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 1 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6454
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="full length human CLASP-4 cDNA"
     CDS             95. .6121
                     /note="human CLASP-4"
                     /codon_start=1
                     /protein_id="CAC43729.1"
                     /db_xref="GI:14597976"
                     /translation="MFPMEDISISVIGRQRRTVQSTVPEDAEKRAQSLFVKECIKTYS
                     TDWHVVNYKYEDFSGDFRMLPCKSLRPEKIPNHVFEIDEDCEKDEDSSSLCSQKGGVI
                     KQGWLHKANVNSTITVTMKVFKRRYFYLTQLPDGSYILNSYKDEKNSKESKGCIYLDA
                     CIDVVQCPKMRRHAFELKMLDKYSHYLAAETEQEMEEWLITLKKIIQINTDSLVQEKK
                     ETVETAQDDETSSQGKAENIMASLERSMHPELMKYGRETEQLNKLSRGDGRQNLFSFD
                     SEVQRLDFSGIEPDIKPFEEKCNKRFLVNCHDLTFNILGQIGDNAKGPPTNVEPFFIN
                     LALFDVKNNCKISADFHVDLNPPSVREMLWGSSTQLASDGSPKGSSPESYIHGIAESQ
                     LRYIQQGIFSVTNPHPEIFLVARIEKVLQGNITHCAEPYIKNSDPVKTAQKVHRTAKQ
                     VCSRLGQYRMPFAWAARPIFKDTQGSLDLDGRFSPLYKQDSSKLSSEDILKLLSEYKK
                     PEKTKLQIIPGQLNITVECVPVDLSNCITSSYVPLKPFEKNCQNITVEVEEFVPEMTK
                     YCYPFTIYKNHLYVYPLQLKYDSQKTFAKARNIAVCVEFRDSDESDASALKCIYGKPA
                     GSVFTTNAYAVVSHHNQNPEFYDEIKIELPIHLHQKHHLLFTFYHVSCEINTKGTTKK
                     QDTVETPVGFAWVPLLKDGRIITFEQQLPVSANLPPGYLNLNDAESRRQCNVDIKWVD
                     GAKPLLKFKSHLESTIYTQDLHVHKFFHHCQLIQSGSKEVPGELIKYLKCLHAMEIQV
                     MIQFLPVILMQLFRVLTNMTHEDDVPINCTMVLLHIVSKCHEEGLDSYLRSFIKYSFR
                     PEKPSAPQAQLIHETLATTMIAILKQSADFLSINKLLKYSWFFFEIIAKSMATYLLEE
                     NKIKLPRGQRFPETYHHVLHSLLLAIIPHVTIRYAEIPDESRNVNYSLASFLKRCLTL
                     MDRGFIFNLINDYISGFSPKDPKVLAEYKFEFLQTICNHEHYIPLNLPMAFAKPKLQR
                     VQDSNLEYSLSDEYCKHHFLVGLLLRETSIALQDNYEIRYTAISVIKNLLIKHAFDTR
                     YQHKNQQAKIAQLYLPFVGLLLENIQRLAGRDTLYSCAAMPNSASRDEFPCGFTSPAN
                     RGSLSTDKDTAYGSFQNGHGIKREDSRGSLIPEGATGFPDQGNTGENTRQSSTRSSVS
                     QYNRLDQYEIRSLLMCYLYIVKMISEDTLLTYWNKVSPQELINILILLEVCLFHFRYM
                     GKRNIARVHDAWLSKHFGIDRKSQTMPALRNRSGVMQARLQHLSSLESSFTLNHSSTT
                     TEADIFHQALLEGNTATEVSLTVLDTISFFTQCFKTQLLNNDGHNPLMKKVFDIHLAF
                     LKNGQSEVSLKHVFASLRAFISKFPSAFFKGRVNMCAAFCYEVLKCCTSKISSTRNEA
                     SALLYLLMRNNFEYTKRKTFLRTHLQIIIAVSQLIADVALSGGSRFQESLFIINNFAN
                     SDRPMKATAFPAEVKDLTKRIRTVLMATAQMKEHEKDPEMLIDLQYSLAKSYASTPEL
                     RKTWLDSMAKIHVKNGDFSEAAMCYVHVAALVAEFLHRKKLFPNGCSAFKKITPNIDE
                     EGAMKEDAGMMDVHYSEEVLLELLEQCVDGLWKAERYEIISEISKLIVPIYEKRREFE
                     KLTQVYRTLHGAYTKILEVMHTKKRLLGTFFRVAFYGQSFFEEEDGKEYIYKEPKLTG
                     LSEISLRLVKLYGEKFGTENVKIIQDSDKVNAKELDPKYAHIQVTYVKPYFDDKELTE
                     RKTEFERNHNISRFVFEAPYTLSGKKQGCIEEQCKRRTILTTSNSFPYVKKRIPINCE
                     QQINLKPIDGATDEIKDKTAELQKLCSSTDVDMIQLQLKLQGWVSVQVNAGPLAYARA
                     FLNDSQASKYPPKKVSELKDMFRKFIQACSIALELNERLIKEDQVEYHEGLKSNFRDM
                     VKELSDIIHEQILQEDTMHSPWMSNTLHVFCAISGTSSDRGYGSPRYAEV"
BASE COUNT     2119 a   1234 c   1303 g   1798 t
ORIGIN

  Query Match             40.9%;  Score 204.4;  DB 6;  Length 6454;
  Best Local Similarity   63.4%;  Pred. No. 4.3e-53;
  Matches  313;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           || |  || ||  ||||||| ||||| || |||||||||||||||||| ||| | |||| |
Db    2000 GTCTCGCATCACAACCAAAATCCAGAGTTCTATGATGAGATTAAAATTGAGCTTCCCATT 2059

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           || || ||| |||| || ||  || |  |||| || |  ||||| || ||||| |
Db    2060 CACCTACATCAAAAACATCATTTGCTTTTCACTTTTTATCATGTAAGTTGTGAAATTAAC 2119

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
           |  || ||||  || || |||   ||    |||||||| | |||||| |   ||||| |
Db    2120 ACAAAGGGAACAACCAAAAAGCAAGACACAGTTGAAACTCCAGTTGGGTTTGCCTGGGTA 2179

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | |||||||| || ||  |  | |||   |||||||  | || || || || ||
Db    2180 CCTTTGCTGAAAGATGGTAGAATCATCACATTTGAGCAGCAGCTGCCAGTTTCCGCCAAT 2239

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           |||||  | |||||  |        | ||    |       | ||||| | |     |||
Db    2240 CTTCCCCCAGGCTACTTGAATCTGAATGATGCAGAATCAAGAAGGCAATGTAACGTGGAT 2299

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           |||||||||||||||||  |  |||||  || ||||  ||    |  ||  | |  |||||
Db    2300 ATTAAATGGGTAGATGGTGCAAAGCCTTTGTTGAAGTTTAAAAGCCACTTAGAATCTACC 2359

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 421
             | || ||||| |||| |||| | || || || |||||  | ||  |||    |   | ||
Db    2360 ATTTACACTCAAGATCTGCATGTGCACAAATTCTTCCATCATTGCCAGCTGATTCAGTCA 2419

Qy     422 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 481
           ||  |  |||   |   |   || ||  | || ||  | ||| || |||||||  |||||
Db    2420 GGCTCGAAAGAAGTTCCAGGGGAGCTCATTAAATATTTAAAGTGTTTGCATGCCATGGAG 2479

Qy     482 GGCCACGTGATGAT 495
             |||  || |||||
Db    2480 ATCCAAGTCATGAT 2493


RESULT 5
CNS01RGX
LOCUS       CNS01RGX               139887 bp    DNA     linear   PRI 26-APR-2001
DEFINITION  Human chromosome 14 DNA sequence BAC C-2373J19 of library CalTech-D
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL160233
VERSION     AL160233.3  GI:13016589
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 139887)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 139887)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Feb 20, 2001 this sequence version replaced gi:7799784.
            --------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : R-857B24 (AC=AL049870)
            Downstream BAC (overlapping the SP6 end) : R-398E10 ----------
            Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 8.87x in Q20 bases; sum-of-contigs
            --------------
```

```
            ----------------------
            Overall quality chart :
            Range     :   bases
            0         :
            1   -  9  :       2
            10  - 19  :       2
            20  - 29  :       3
            30  - 39  :      28
            40  - 49  :     816
            50  - 59  :    1974
            60  - 69  :    2026
            70  - 79  :    6616
            80  - 89  :   31804
            90  - 99  :   96616
            ----------------------
            Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1. .139887
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="C-2373J19"
                     /clone_lib="CalTech-D"
BASE COUNT    41045 a  25978 c  27103 g  45761 t
ORIGIN

  Query Match             39.0%;  Score 194.8;  DB 9;  Length 139887;
  Best Local Similarity   62.1%;  Pred. No. 8.4e-50;
  Matches  307;  Conservative    0;  Mismatches  187;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
            ||  |  |||  ||   |  |||||||||||  ||  ||||||||||||||||||  |||  |  |||   |
Db     421 GTCTCACATCACAATCAAAACCCAGAGTTCTATGATGAGATTAAAATTGAGCTTCCCTTT 480

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
            ||  |||||  ||||  ||  ||  ||  |  ||||  ||  |  |||  |  ||  ||||| |
Db     481 CACTTGCATCAAAAACATCATTTGCTTTTCACTTTTTGTCATATAAGTTGTGAAATTAAC 540

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
            |  ||  ||||  ||  ||  |||   ||    ||||||||  |  |||||  |   |||||  |
Db     541 ACAAAGGGAACAACCAAAAAGCAAGACACAGTTGAAACTCCAGTTGAGTTTGCCTGGGTA 600

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
            ||  |  ||||||||  ||  ||  |  |  |||   |  ||||||  |  ||  ||  ||  |  ||
Db     601 CCATTGCTGAAAGATGGTAGAATCATCACATTTGGGCAGCAGCTGCCAGTTTCCACCAAT 660

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
            |||||  |  |||||  |        |  ||           |  |||||  |  |   |  |||
Db     661 CTTCCCCCAGGCTACTTAAATCTGAATGATGGATAATCAAGAAGGCAATGTAATGTGGAT 720

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
            ||||||||||||||||  |  |||||  ||  |||  |||  |  ||  |  |  |||||
Db     721 ATTAAATGGGTAGATGGTGTAAAGCCTTTGTGGAAGATTAAAAGCCACTTAGAATCTACC 780

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 421
            |  ||  |||  |  ||||  ||||||  ||  ||  ||  |||||  |  ||  |||  |  |  ||
Db     781 ATTTACACTTAAGATCTGCATTTGCACAAATTCTTCCATCATTGCCAGCTGAATCAGTCA 840

Qy     422 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 481
            |  |  |||  |  |  ||  ||  |  ||  ||  |  |||  |  |||||||  |||||
Db     841 GACTCGAAAGAAGTTCCAGGGGAGCTCATTAAATATTTAAAGTCTTTGCATGCCATGGAG 900

Qy     482 GGCCACGTGATGAT 495
              |||  ||  |||||
Db     901 ATCCAAGTCATGAT 914


RESULT 6
AL161420/c
LOCUS       AL161420              163316 bp    DNA     linear   PRI 01-FEB-2001
DEFINITION  Human DNA sequence from clone RP11-155N3 on chromosome 13 Contains
            ESTs, STSs and GSSs. Contains the 3' part of a novel gene similar
            to KIAA0694, the KIAA1058 gene and a putative novel gene, complete
            sequence.
ACCESSION   AL161420
VERSION     AL161420.10  GI:10443397
KEYWORDS    HTG; KIAA0694; KIAA1058.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163316)
  AUTHORS   Smith,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 1, 2000 this sequence version replaced gi:10039689.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            This sequence is the entire insert of clone RP11-155N3 The true
            left end of clone RP11-318G11 is at 114983 in this sequence. The
            true right end of clone RP11-56D6 is at 42341 in this sequence. The
            true right end of clone RP11-551M18 is at 43464 in this sequence.
            This sequence has been finished according to sequence map criteria
            as follows.  An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated repeat sequence elements.  Where the sequence is
            ambiguous, there is an annotation using the 'unsure' feature key.
            RP11-155N3 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .163316
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-155N3"
                     /clone_lib="RPCI-11.1"
     misc_feature    6. .510
                     /note="match: GSS: Em:AQ372034"
     repeat_region   75. .372
                     /note="AluJb repeat: matches 1. .298 of consensus"
     repeat_region   447. .742
                     /note="AluSp repeat: matches 1. .296 of consensus"
     repeat_region   745. .1025
                     /note="L1MC1 repeat: matches 6019. .6327 of consensus"
     repeat_region   1250. .1566
                     /note="AluSx repeat: matches 2. .311 of consensus"
     repeat_region   1631. .2325
                     /note="L1MC5 repeat: matches 6683. .7400 of consensus"
     repeat_region   2326. .2619
                     /note="AluSx repeat: matches 1. .295 of consensus"
     repeat_region   2620. .2890
                     /note="L1MC5 repeat: matches 7400. .7665 of consensus"
     misc_feature    complement(3019. .3311)
                     /note="match: STS: Em:HSA273ZH1"
     misc_feature    3060. .3191
                     /note="match: GSS: Em:B17811"
     misc_feature    complement(3065. .3246)
                     /note="match: STS: Em:AF009212"
     misc_feature    3067. .3190
```

```
                     /note="match: GSS: Em:AQ114323"
     repeat_region   3087. .3122
                     /note="18 copies 2 mer tg 97% conserved"
     repeat_region   3089. .3124
                     /note="6 copies 6 mer tgtgtg 97% conserved"
     misc_feature    3123. .3280
                     /note="match: GSS: Em:AQ459714"
     repeat_region   3191. .3475
                     /note="AluSx repeat: matches 1. .288 of consensus"
     repeat_region   3692. .3773
                     /note="MIR repeat: matches 3. .94 of consensus"
     repeat_region   3887. .4074
                     /note="L2 repeat: matches 1607. .1787 of consensus"
     repeat_region   4082. .4144
                     /note="3 copies 21 mer 88% conserved"
     repeat_region   4418. .4630
                     /note="L2 repeat: matches 1354. .1587 of consensus"
     repeat_region   7208. .7609
                     /note="MLT1B repeat: matches 8. .390 of consensus"
     repeat_region   9220. .9344
                     /note="MIR repeat: matches 71. .196 of consensus"
     repeat_region   9586. .9667
                     /note="L1ME repeat: matches 5569. .5651 of consensus"
     repeat_region   9714. .9813
                     /note="L2 repeat: matches 771. .874 of consensus"
     repeat_region   9869. .10202
                     /note="L2 repeat: matches 971. .1261 of consensus"
     repeat_region   10203. .10581
                     /note="MSTA repeat: matches 1. .426 of consensus"
     repeat_region   10582. .11267
                     /note="L2 repeat: matches 1261. .2121 of consensus"
     repeat_region   11460. .11570
                     /note="L2 repeat: matches 2377. .2497 of consensus"
     repeat_region   11603. .11778
                     /note="L1MB3 repeat: matches 5996. .6166 of consensus"
     repeat_region   11779. .12160
                     /note="MER51A repeat: matches 1. .371 of consensus"
     repeat_region   12161. .12729
                     /note="L1MB3 repeat: matches 5403. .5996 of consensus"
     repeat_region   12682. .12950
                     /note="L1MD repeat: matches -1. .274 of consensus"
     repeat_region   12945. .13209
                     /note="L1MC5 repeat: matches 7249. .7517 of consensus"
     repeat_region   13323. .13413
                     /note="AluJ/FRAM repeat: matches 221. .311 of consensus"
     repeat_region   13711. .13846
                     /note="AluSq/x repeat: matches 1. .136 of consensus"
     misc_feature    13955. .14417
                     /note="match: GSS: Em:AQ475161"
     misc_feature    complement(14122. .14590)
                     /note="match: GSS: Em:AQ349540"
     repeat_region   14334. .14400
                     /note="L1MA10 repeat: matches 6204. .6272 of consensus"
     repeat_region   14407. .14693
                     /note="AluSx repeat: matches 1. .288 of consensus"
     misc_feature    14584. .15413
                     /note="match: GSS: Em:AQ739876"
     repeat_region   14719. .14793
                     /note="L1MC/D repeat: matches 5744. .5818 of consensus"
     repeat_region   15013. .15193
                     /note="L1MB3 repeat: matches 5993. .6184 of consensus"
     gene            complement(join(16480. .17741,19199. .19206,20108. .20246,
                     20382. .20482,23281. .23471,27992. .28202,30654. .30800,
                     31633. .31668,32116. .32121,32345. .32458,33162. .33284,
                     47391. .47497,48865. .48933,49823. .49880,52300. .52515,
                     52636. .52749,54341. .54523,54642. .54779,60487. .60602,
                     68322. .68365,68917. .69053,71585. .71632,73040. .73106,
                     76401. .76515,78891. .79024,83174. .83515,86012. .86125,
                     86363. .86541,90418. .90529,90921. .91004,91224. .91370,
                     102891. .102962,103540. .103665,104551. .104620,
                     104885. .105000,106035. .106124,106806. .106892,
                     107964. .108085,108701. .108824,109525. .109617,
                     111151. .111216,111353. .111540,111625. .111704,
                     120481. .120603))
                     /gene="bA155N3.2"
     mRNA            complement(join(16480. .17741,19199. .19206,20108. .20246,
                     20382. .20482,23281. .23471,27992. .28202,30654. .30800,
                     31633. .31668,32116. .32121,32345. .32458,33162. .33284,
                     47391. .47497,48865. .48933,49823. .49880,52300. .52515,
                     52636. .52749,54341. .54523,54642. .54779,60487. .60602,
                     68917. .68984,71585. .71632,73040. .73106,76401. .76515,
                     78891. .79024,83398. .83515,86012. .86125,86363. .86541,
                     90418. .90529,90921. .91004,91224. .91370,102891. .102962,
                     103540. .103665,104551. .104620,104885. .105000,
                     106035. .106124,106806. .106892,107964. .108085,
                     108701. .108824,109525. .109617,111151. .111216,
                     111353. .111540,111625. .111704,120481. .>120603))
                     /gene="bA155N3.2"
                     /product="bA155N3.2.1 (KIAA1058)"
                     /note="match: cDNAs: Em:AB028981 Em:U79289
                     match: ESTs: Em:W27872 Em:H29845 Em:H29936 Em:AW160928
                     Em:AW457794 Em:T62004 Em:R64450 Em:D58813 Em:H10596
                     Em:AW380042 Em:AW344638 Em:AW450751 Em:AF188523
                     Em:AA151055 Em:T33377 Em:AW466130 Em:AI014329 Em:Z42960
                     Em:AI372817 Em:AI372818 Em:AW389763 Em:AW230012
                     Em:AL133724 Em:AW162535 Em:AW607903 Em:AA484945 Em:H28862
                     Em:Z39079 Em:AI767696 Em:AW576807 Em:AI078244 Em:R78583
                     Em:AA705585 Em:AI271453 Em:AW512964 Em:W88550 Em:AI870702
                     Em:AL079478 Em:AI693777 Em:AW576805 Em:AI467994
                     Em:AI372819 Em:AI732576 Em:AI277278 Em:AI969485
                     Em:AI768094 Em:AA779769 Em:AA502043 Em:R27124 Em:AW576813
                     Em:AW576816 Em:AW168032 Em:AI365152 Em:AA165512
                     Em:AI120344 Em:AI888868 Em:AA165554 Em:R73986 Em:AA523853
                     Em:AI041815 Em:AL134605 Em:AW009821 Em:D62808 Em:AI769357
                     Em:AW363093 Em:AI733873 Em:W88496 Em:T03019 Em:AW071481
                     Em:AI034000"
                     /evidence=not_experimental
     polyA_site      complement(16480)
                     /gene="bA155N3.2"
     polyA_site      complement(16482)
                     /gene="bA155N3.2"
     polyA_site      complement(16483)
                     /gene="bA155N3.2"
     polyA_signal    complement(16498. .16503)
                     /gene="bA155N3.2"
     misc_feature    16511. .16637
                     /note="match: STS: Em:G42960"
     polyA_signal    complement(16523. .16528)
                     /gene="bA155N3.2"
     polyA_signal    complement(16570. .16575)
                     /gene="bA155N3.2"
     CDS             complement(join(17741,19199. .19206,20108. .20246,
                     20382. .20482,23281. .23471,27992. .28202,30654. .30800,

  Query Match             25.4%;  Score 126.8;  DB 9;  Length 163316;
  Best Local Similarity   98.5%;  Pred. No. 2.9e-28;
  Matches  128;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      39 AGATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCT 98
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108826 AGATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCT 108767

Qy      99 TCCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAA 158
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108766 TCCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAA 108707

Qy     159 CCCAAGTTGG 168
           |||||| | |
Db 108706 CCCAAGGTTG 108697


RESULT 7
AC109966/c
LOCUS       AC109966               175281 bp    DNA     linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-323I5, *** SEQUENCING IN PROGRESS
            ***, 75 unordered pieces.
```

```
ACCESSION   AC109966
VERSION     AC109966.3  GI:21738210
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 175281)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 12, 2002 this sequence version replaced gi:18847026.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GQLP
                Center clone name: CH230-323I5
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 105442 bases at least Q40
                Consensus quality: 111492 bases at least Q30
                Consensus quality: 116270 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 75 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1290: contig of 1290 bp in length
            *     1291     1390: gap of unknown length
            *     1391     2414: contig of 1024 bp in length
            *     2415     2514: gap of unknown length
            *     2515     3639: contig of 1125 bp in length
            *     3640     3739: gap of unknown length
            *     3740     5063: contig of 1324 bp in length
            *     5064     5163: gap of unknown length
            *     5164     6243: contig of 1080 bp in length
            *     6244     6343: gap of unknown length
            *     6344     7536: contig of 1193 bp in length
            *     7537     7636: gap of unknown length
            *     7637     9071: contig of 1435 bp in length
            *     9072     9171: gap of unknown length
            *     9172    10200: contig of 1029 bp in length
            *    10201    10300: gap of unknown length
            *    10301    11505: contig of 1205 bp in length
            *    11506    11605: gap of unknown length
            *    11606    12680: contig of 1075 bp in length
            *    12681    12780: gap of unknown length
            *    12781    14056: contig of 1276 bp in length
            *    14057    14156: gap of unknown length
            *    14157    15413: contig of 1257 bp in length
            *    15414    15513: gap of unknown length
            *    15514    16580: contig of 1067 bp in length
            *    16581    16680: gap of unknown length
            *    16681    17692: contig of 1012 bp in length
            *    17693    17792: gap of unknown length
            *    17793    19013: contig of 1221 bp in length
            *    19014    19113: gap of unknown length
            *    19114    20418: contig of 1305 bp in length
            *    20419    20518: gap of unknown length
            *    20519    21880: contig of 1362 bp in length
            *    21881    21980: gap of unknown length
            *    21981    22985: contig of 1005 bp in length
            *    22986    23085: gap of unknown length
            *    23086    24156: contig of 1071 bp in length
            *    24157    24256: gap of unknown length
            *    24257    25732: contig of 1476 bp in length
            *    25733    25832: gap of unknown length
            *    25833    27196: contig of 1364 bp in length
            *    27197    27296: gap of unknown length
            *    27297    28615: contig of 1319 bp in length
            *    28616    28715: gap of unknown length
            *    28716    30145: contig of 1430 bp in length
            *    30146    30245: gap of unknown length
            *    30246    31418: contig of 1173 bp in length
            *    31419    31518: gap of unknown length
            *    31519    32896: contig of 1378 bp in length
            *    32897    32996: gap of unknown length
            *    32997    34142: contig of 1146 bp in length
            *    34143    34242: gap of unknown length
            *    34243    35262: contig of 1020 bp in length
            *    35263    35362: gap of unknown length
            *    35363    36924: contig of 1562 bp in length
            *    36925    37024: gap of unknown length
```

```
  *    37025    38670: contig of 1646 bp in length
  *    38671    38770: gap of unknown length
  *    38771    40612: contig of 1842 bp in length
  *    40613    40712: gap of unknown length
  *    40713    42724: contig of 2012 bp in length
  *    42725    42824: gap of unknown length
  *    42825    43829: contig of 1005 bp in length
  *    43830    43929: gap of unknown length
  *    43930    45818: contig of 1889 bp in length
  *    45819    45918: gap of unknown length
  *    45919    48183: contig of 2265 bp in length
  *    48184    48283: gap of unknown length
  *    48284    49887: contig of 1604 bp in length
  *    49888    49987: gap of unknown length
  *    49988    51568: contig of 1581 bp in length
  *    51569    51668: gap of unknown length
  *    51669    53251: contig of 1583 bp in length
  *    53252    53351: gap of unknown length
  *    53352    55636: contig of 2285 bp in length
  *    55637    55736: gap of unknown length
  *    55737    57159: contig of 1423 bp in length
  *    57160    57259: gap of unknown length
  *    57260    59519: contig of 2260 bp in length
  *    59520    59619: gap of unknown length
  *    59620    63392: contig of 3773 bp in length
  *    63393    63492: gap of unknown length
  *    63493    65130: contig of 1638 bp in length
  *    65131    65230: gap of unknown length
  *    65231    68145: contig of 2915 bp in length
  *    68146    68245: gap of unknown length
  *    68246    70070: contig of 1825 bp in length
  *    70071    70170: gap of unknown length
  *    70171    73494: contig of 3324 bp in length
  *    73495    73594: gap of unknown length
  *    73595    75991: contig of 2397 bp in length
  *    75992    76091: gap of unknown length
  *    76092    77867: contig of 1776 bp in length
  *    77868    77967: gap of unknown length
  *    77968    80318: contig of 2351 bp in length
  *    80319    80418: gap of unknown length
  *    80419    82211: contig of 1793 bp in length
  *    82212    82311: gap of unknown length
  *    82312    84412: contig of 2101 bp in length
  *    84413    84512: gap of unknown length
  *    84513    85982: contig of 1470 bp in length
  *    85983    86082: gap of unknown length
  *    86083    88348: contig of 2266 bp in length
  *    88349    88448: gap of unknown length
  *    88449    90746: contig of 2298 bp in length
  *    90747    90846: gap of unknown length
  *    90847    92600: contig of 1754 bp in length

  Query Match             18.1%;  Score 90.4;  DB 2;  Length 175281;
  Best Local Similarity   83.1%;  Pred. No. 9.8e-17;
  Matches  103;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy     163 AGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCA 222
           |||||||| |||||||||||| ||||||||||||||| || ||| |||| |  || |||||
Db   84981 AGTTGGCTTCTCCTGGCTTCCTCTCCTGAAAGACGGCAGAGTGTTGACGAATGAACAGCA 84922

Qy     223 CATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGG 282
           ||||| ||||| || ||||| || || ||||| |||||||||| ||||| | |||||
Db   84921 TATCCCTGTCTCCGCCAACCTGCCGTCTGGCTACCTTGGCTACCAGGAGCTCAGCATGGG 84862

Qy     283 CAGG 286
           ||||
Db   84861 CAGG 84858


RESULT 8
AX173028
LOCUS       AX173028               4391 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 7 from Patent WO0142294.
ACCESSION   AX173028
VERSION     AX173028.1  GI:14597978
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4391)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 7 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .4391
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="human CLASP-4 cDNA"
     CDS             414. .4058
                     /note="human CLASP-4"
                     /codon_start=1
                     /protein_id="CAC43730.1"
                     /db_xref="GI:14597979"
                     /translation="MEIQVLIRFLSVILMQLFWVLPNMIHEDDVPISCPMVLFHIVSK
                     CHEEGLDSYLSSFIKYSFRPGKPSAPQAPLIHETLATMMIALLKQSADFLAINKLLKY
                     SWFFFEIIAKSMATYLLEENKIKLTHGQRFPKAYHHALHSLFLAITIVESQYAEIPKE
                     SRNVNYSLASFLKCCLTLMDRGFVFNLINDYISGFSPKDPKVLAEYKFEFLQTICNHE
                     HYIPLNLPMAFAKPKLQRVQDSNLEYSLSDEYCKHHFLVGLLLRETSIALQDNYEIRY
                     TAISVIKNLLIKHAFDTRYQHKNQQAKIAQLYLPFVGLLLENIQRLAGRDTLYSCAAM
                     PNSASRDEFPCGFTSPANRGSLSTDKDTAYGSFQNGHGIKREDSRGSLIPEGATGFPD
                     QGNTGENTRQSSTRSSVSQYNRLDQYEIRSLLMCYLYIVKMISEDTLLTYWNKVSPQE
                     LINILILLEVCLFHFRYMGKRNIARVHDAWLSKHFGIDRKSQTMPALRNRSGVMQARL
                     QHLSSLESSFTLNHSSTTTEADIFHQALLEGNTATEVSLTVLDTISFFTQCFKTHFLN
                     NDGHNPLMKKVFDIHLAFLKNGQSEVSLKHVFASLRAFISKFPSAFFKGRVNMCAAFC
                     YEVLKCCTSKISSTRNEASALLYLLMRNNFEYTKRKTFLRTHLQIIIAVSQLIADVAL
                     SGGSRFQESLFIINNFANSDRPMLARAFPAEVKDLTKRIRTVLMATAQMKEHEKDPEM
                     LIDLQYSLAKSYASTPELRKTWLDSMAKIHVKNGDFSEAAMCYVHVAALVAEFLHRKK
                     LFPNGCSAFKKITPNIDEEGAMKEDAGMMDVHYSEEVLLELLEQCVNGLWKAERYEII
                     SEISKLIGPIYENRREFENLTQVYRTLHGAYTKILEVMHTKKRLLGTFFRVAFYGQSF
                     FEEEDGKEYIYKEPKLTGLSEISLRLVKLYGEKFGTENVKIIQDSDKVNAKELDPKYA
                     HIQVTYVKPYFDDKELTERKTEFERNHNISRFVFEAPYTLSGKKQGCIEEQCKRRTIL
                     TTSNSFPYVKKRIPINCEQQINLKPIDGATDEIKDKTAELQKLCSSTDVDMIQLQLKL
                     QGWVSVQVNAGPLAYARAFLNDSQASKYPPKKVSELKDMFRKFIQACSIALELNERLI
                     KEDQVEYHEGLKSNFRDMVKELSDIIHEQILQEDTMHSPWMSNTLHVFCAISGTSSDR
                     GYGSPRYAEV"
BASE COUNT     1359 a    851 c    908 g   1273 t
ORIGIN

  Query Match             17.0%;  Score 85;  DB 6;  Length 4391;
  Best Local Similarity   52.4%;  Pred. No. 2.5e-15;
  Matches  187;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy     140 AAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGA 199
           |||  ||    || |  | | |  ||||| |   ||||| | ||  | ||||| |  ||
Db      78 AAGGCGGGCCCCGGTTGACCTCCGGTTGGGTTTGCCTGGGTCCCTTTGCTGAAGGTTGGT 137

Qy     200 AGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTT 259
           ||  |  | |    | || ||   | ||||| ||  | |  |||||  ||||||   |
Db     138 AGATTCTTCCCTTTTGGGCGGCGGCTGCCGGTTTCCCCCATTCTTCCCCCGGGCTCCTTA 197

Qy     260 GGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGA 319
                  ||          |  |||| | |  |  |||  ||||||||| | |||
Db     198 ATTCTGATTGATGGATAATCAAGGGGGCAATGTAATGTGGATTTTAAATGGGTGGTTGGT 257

Qy     320 GGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAG 379
           |  |||||  ||  ||   ||      |   ||| |||| |  | |   || ||| || |
Db     258 GTAAAGCCTTTGTGGAGGTTTAAAGGCCCCTTGGGTTCTCCCATTTCCCCTTAGGTTCTG 317

Qy     380 CATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGA 439
           ||||| |  || || |||||  | || | |    |     || |   |  |||   |  |
Db     318 CATTTGCCCAAATTCTTCCATCATTGCCGGCTGATTCGGTCAGACTCGAAAGAGGTTCCA 377

Qy     440 AACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAAGGCCACGTGATGATC 496
             |  ||  | || |   | | |  | |||||| |||||   ||| ||  |||||
```

```
Db    378 GGGGGGCTCATTAAATTTTTAAGGTCTTTGCATGCCATGGAGATCCAAGTCTTGATC 434


RESULT 9
AX173118
LOCUS       AX173118               4393 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 97 from Patent WO0142294.
ACCESSION   AX173118
VERSION     AX173118.1  GI:14598012
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4393)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 97 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .4393
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="polynucleotide fragment"
BASE COUNT     1359 a    852 c    910 g   1272 t
ORIGIN

  Query Match             17.0%;  Score 85;  DB 6;  Length 4393;
  Best Local Similarity   52.4%;  Pred. No. 2.5e-15;
  Matches  187;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy    140 AAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGA 199
          |||  ||    || |  | | |  ||||| |   ||||| | ||  | ||||| |  ||
Db     78 AAGGCGGGCCCCCGGTTGACCTCCGGTTGGGTTTGCCTGGGTCCCTTTGCTGAAGGTTGGT 137

Qy    200 AGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTT 259
          ||  |  |  |    | || ||   | ||||| ||  | |  |||||  ||||||   |
Db    138 AGATTCTTCCCTTTTGGGCGGCGGCTGCCGGTTTCCCCCATTCTTCCCCCGGGCTCCTTA 197

Qy    260 GGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGA 319
                 ||          | |||| | |  |  |||  |||||||||| | |||
Db    198 ATTCTGATTGATGGATAATCAAGGGGGCAATGTAATGTGGATTTTAAATGGGTGGTTGGT 257

Qy    320 GGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAG 379
          |  |||||  ||  ||   ||      |   ||| |||| |  | |   || ||| || |
Db    258 GTAAAGCCTTTGTGGAGGTTTAAAGGCCCCTTGGGTTCTCCCATTTCCCCTTAGGTTCTG 317

Qy    380 CATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGA 439
          |||||  |  || || |||||  | || | |    |     || |   |  |||   |   |
Db    318 CATTTGCCCAAATTCTTCCATCATTGCCGGCTGATTCGGTCAGACTCGAAAGAGGTTCCA 377

Qy    440 AACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAAGGCCACGTGATGATC 496
            |  ||  | || |   | | |  | |||||| |||||   ||| ||  |||||
Db    378 GGGGGGCTCATTAAATTTTTAAGGTCTTTGCATGCCATGGAGATCCAAGTCTTGATC 434


RESULT 10
AC109966
LOCUS       AC109966               175281 bp    DNA     linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-323I5, *** SEQUENCING IN PROGRESS
            ***, 75 unordered pieces.
ACCESSION   AC109966
VERSION     AC109966.3  GI:21738210
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 175281)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 12, 2002 this sequence version replaced gi:18847026.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GQLP
                Center clone name: CH230-323I5
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 105442 bases at least Q40
                Consensus quality: 111492 bases at least Q30
                Consensus quality: 116270 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 75 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
```

```
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1290: contig of 1290 bp in length
*     1291     1390: gap of unknown length
*     1391     2414: contig of 1024 bp in length
*     2415     2514: gap of unknown length
*     2515     3639: contig of 1125 bp in length
*     3640     3739: gap of unknown length
*     3740     5063: contig of 1324 bp in length
*     5064     5163: gap of unknown length
*     5164     6243: contig of 1080 bp in length
*     6244     6343: gap of unknown length
*     6344     7536: contig of 1193 bp in length
*     7537     7636: gap of unknown length
*     7637     9071: contig of 1435 bp in length
*     9072     9171: gap of unknown length
*     9172    10200: contig of 1029 bp in length
*    10201    10300: gap of unknown length
*    10301    11505: contig of 1205 bp in length
*    11506    11605: gap of unknown length
*    11606    12680: contig of 1075 bp in length
*    12681    12780: gap of unknown length
*    12781    14056: contig of 1276 bp in length
*    14057    14156: gap of unknown length
*    14157    15413: contig of 1257 bp in length
*    15414    15513: gap of unknown length
*    15514    16580: contig of 1067 bp in length
*    16581    16680: gap of unknown length
*    16681    17692: contig of 1012 bp in length
*    17693    17792: gap of unknown length
*    17793    19013: contig of 1221 bp in length
*    19014    19113: gap of unknown length
*    19114    20418: contig of 1305 bp in length
*    20419    20518: gap of unknown length
*    20519    21880: contig of 1362 bp in length
*    21881    21980: gap of unknown length
*    21981    22985: contig of 1005 bp in length
*    22986    23085: gap of unknown length
*    23086    24156: contig of 1071 bp in length
*    24157    24256: gap of unknown length
*    24257    25732: contig of 1476 bp in length
*    25733    25832: gap of unknown length
*    25833    27196: contig of 1364 bp in length
*    27197    27296: gap of unknown length
*    27297    28615: contig of 1319 bp in length
*    28616    28715: gap of unknown length
*    28716    30145: contig of 1430 bp in length
*    30146    30245: gap of unknown length
*    30246    31418: contig of 1173 bp in length
*    31419    31518: gap of unknown length
*    31519    32896: contig of 1378 bp in length
*    32897    32996: gap of unknown length
*    32997    34142: contig of 1146 bp in length
*    34143    34242: gap of unknown length
*    34243    35262: contig of 1020 bp in length
*    35263    35362: gap of unknown length
*    35363    36924: contig of 1562 bp in length
*    36925    37024: gap of unknown length
*    37025    38670: contig of 1646 bp in length
*    38671    38770: gap of unknown length
*    38771    40612: contig of 1842 bp in length
*    40613    40712: gap of unknown length
*    40713    42724: contig of 2012 bp in length
*    42725    42824: gap of unknown length
*    42825    43829: contig of 1005 bp in length
*    43830    43929: gap of unknown length
*    43930    45818: contig of 1889 bp in length
*    45819    45918: gap of unknown length
*    45919    48183: contig of 2265 bp in length
*    48184    48283: gap of unknown length
*    48284    49887: contig of 1604 bp in length
*    49888    49987: gap of unknown length
*    49988    51568: contig of 1581 bp in length
*    51569    51668: gap of unknown length
*    51669    53251: contig of 1583 bp in length
*    53252    53351: gap of unknown length
*    53352    55636: contig of 2285 bp in length
*    55637    55736: gap of unknown length
*    55737    57159: contig of 1423 bp in length
*    57160    57259: gap of unknown length
*    57260    59519: contig of 2260 bp in length
*    59520    59619: gap of unknown length
*    59620    63392: contig of 3773 bp in length
*    63393    63492: gap of unknown length
*    63493    65130: contig of 1638 bp in length
*    65131    65230: gap of unknown length
*    65231    68145: contig of 2915 bp in length
*    68146    68245: gap of unknown length
*    68246    70070: contig of 1825 bp in length
*    70071    70170: gap of unknown length
*    70171    73494: contig of 3324 bp in length
*    73495    73594: gap of unknown length
*    73595    75991: contig of 2397 bp in length
*    75992    76091: gap of unknown length
*    76092    77867: contig of 1776 bp in length
*    77868    77967: gap of unknown length
*    77968    80318: contig of 2351 bp in length
*    80319    80418: gap of unknown length
*    80419    82211: contig of 1793 bp in length
*    82212    82311: gap of unknown length
*    82312    84412: contig of 2101 bp in length
*    84413    84512: gap of unknown length
*    84513    85982: contig of 1470 bp in length
*    85983    86082: gap of unknown length
*    86083    88348: contig of 2266 bp in length
*    88349    88448: gap of unknown length
*    88449    90746: contig of 2298 bp in length
*    90747    90846: gap of unknown length
*    90847    92600: contig of 1754 bp in length

  Query Match             15.2%;  Score 75.8;  DB 2;  Length 175281;
  Best Local Similarity   68.9%;  Pred. No. 4.1e-12;
  Matches  104;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy     323 AAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCAT 382
           ||||     || | |  ||    ||  ||  |  |       |  ||  ||||||||||
Db   83560 AAGCACACACTAACAGATTTTCTTCTCCTTCTCCCCCTGCCCTCAACCTAGGATCAGCAC 83619

Qy     383 TTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAAC 442
           |||||  |||||||||||  |||||||||||||  |||||||||||||||||||  |||  |
Db   83620 TTACACAATTTTTTCCAATACTGTCAGAAAACGGAATCTGGAGCCCAAGCCTCAGGGAGT 83679

Qy     443 GAACTTGTAAAGTACCTTAAGAGTCTGCATG 473
           |||||  ||  ||  |||||||||   |   |||
Db   83680 GAACTAGTGAAATACCTTAAGGTACCAAATG 83710


RESULT 11
AC119357/c
LOCUS       AC119357               192825 bp    DNA     linear   HTG 18-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-473M19, *** SEQUENCING IN PROGRESS
            ***, 89 unordered pieces.
ACCESSION   AC119357
VERSION     AC119357.3  GI:21746916
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 192825)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
```

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 192825)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 192825)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 14, 2002 this sequence version replaced gi:20429771.

---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GVFP
Center clone name: CH230-473M19
---------- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 117932 bases at least Q40
Consensus quality: 124308 bases at least Q30
Consensus quality: 129203 bases at least Q20
----

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 89 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	1129: contig of 1129 bp in length
*	1130	1229: gap of unknown length
*	1230	2849: contig of 1620 bp in length
*	2850	2949: gap of unknown length
*	2950	4125: contig of 1176 bp in length
*	4126	4225: gap of unknown length
*	4226	5644: contig of 1419 bp in length
*	5645	5744: gap of unknown length
*	5745	6997: contig of 1253 bp in length
*	6998	7097: gap of unknown length
*	7098	8757: contig of 1660 bp in length
*	8758	8857: gap of unknown length
*	8858	9860: contig of 1003 bp in length
*	9861	9960: gap of unknown length
*	9961	11605: contig of 1645 bp in length
*	11606	11705: gap of unknown length
*	11706	12751: contig of 1046 bp in length
*	12752	12851: gap of unknown length
*	12852	14490: contig of 1639 bp in length
*	14491	14590: gap of unknown length
*	14591	16101: contig of 1511 bp in length
*	16102	16201: gap of unknown length
*	16202	17628: contig of 1427 bp in length
*	17629	17728: gap of unknown length
*	17729	18883: contig of 1155 bp in length
*	18884	18983: gap of unknown length
*	18984	20211: contig of 1228 bp in length
*	20212	20311: gap of unknown length
*	20312	21353: contig of 1042 bp in length
*	21354	21453: gap of unknown length
*	21454	22518: contig of 1065 bp in length
*	22519	22618: gap of unknown length
*	22619	24107: contig of 1489 bp in length
*	24108	24207: gap of unknown length
*	24208	25603: contig of 1396 bp in length
*	25604	25703: gap of unknown length
*	25704	27171: contig of 1468 bp in length
*	27172	27271: gap of unknown length
*	27272	28993: contig of 1722 bp in length
*	28994	29093: gap of unknown length
*	29094	30122: contig of 1029 bp in length
*	30123	30222: gap of unknown length
*	30223	31317: contig of 1095 bp in length
*	31318	31417: gap of unknown length
*	31418	33144: contig of 1727 bp in length
*	33145	33244: gap of unknown length
*	33245	34431: contig of 1187 bp in length
*	34432	34531: gap of unknown length
*	34532	35743: contig of 1212 bp in length
*	35744	35843: gap of unknown length
*	35844	37180: contig of 1337 bp in length
*	37181	37280: gap of unknown length
*	37281	38596: contig of 1316 bp in length
*	38597	38696: gap of unknown length
*	38697	40031: contig of 1335 bp in length
*	40032	40131: gap of unknown length
*	40132	41485: contig of 1354 bp in length
*	41486	41585: gap of unknown length
*	41586	42744: contig of 1159 bp in length
*	42745	42844: gap of unknown length
*	42845	45158: contig of 2314 bp in length
*	45159	45258: gap of unknown length
*	45259	46752: contig of 1494 bp in length
*	46753	46852: gap of unknown length
*	46853	48029: contig of 1177 bp in length
*	48030	48129: gap of unknown length
*	48130	49682: contig of 1553 bp in length
*	49683	49782: gap of unknown length

```
  *    49783    51724: contig of 1942 bp in length
  *    51725    51824: gap of unknown length
  *    51825    53137: contig of 1313 bp in length
  *    53138    53237: gap of unknown length
  *    53238    55287: contig of 2050 bp in length
  *    55288    55387: gap of unknown length
  *    55388    57026: contig of 1639 bp in length
  *    57027    57126: gap of unknown length
  *    57127    58750: contig of 1624 bp in length
  *    58751    58850: gap of unknown length
  *    58851    60180: contig of 1330 bp in length
  *    60181    60280: gap of unknown length
  *    60281    61971: contig of 1691 bp in length
  *    61972    62071: gap of unknown length
  *    62072    63893: contig of 1822 bp in length
  *    63894    63993: gap of unknown length
  *    63994    65782: contig of 1789 bp in length
  *    65783    65882: gap of unknown length
  *    65883    67125: contig of 1243 bp in length
  *    67126    67225: gap of unknown length
  *    67226    69485: contig of 2260 bp in length
  *    69486    69585: gap of unknown length
  *    69586    71428: contig of 1843 bp in length
  *    71429    71528: gap of unknown length
  *    71529    73230: contig of 1702 bp in length
  *    73231    73330: gap of unknown length
  *    73331    75374: contig of 2044 bp in length
  *    75375    75474: gap of unknown length
  *    75475    77927: contig of 2453 bp in length
  *    77928    78027: gap of unknown length
  *    78028    79268: contig of 1241 bp in length
  *    79269    79368: gap of unknown length
  *    79369    81963: contig of 2595 bp in length
  *    81964    82063: gap of unknown length
  *    82064    83940: contig of 1877 bp in length
  *    83941    84040: gap of unknown length
  *    84041    85690: contig of 1650 bp in length
  *    85691    85790: gap of unknown length
  *    85791    88775: contig of 2985 bp in length

  Query Match             15.2%;  Score 75.8;  DB 2;  Length 192825;
  Best Local Similarity   68.9%;  Pred. No. 4.2e-12;
  Matches  104;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy     323 AAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCAT 382
           ||||     || | |  ||    ||  ||  |  |       |  ||  ||||||||||
Db   33590 AAGCACACACTAACAGATTTTCTTCTCCTTCTCCCCCTGCCCTCAACCTAGGATCAGCAC 33531

Qy     383 TTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAAC 442
           ||||| |||||||||||| |||||||||||||| ||||||||||||||||||| ||| |
Db   33530 TTACACAATTTTTTCCAATACTGTCAGAAAACGGAATCTGGAGCCCAAGCCTCAGGGAGT 33471

Qy     443 GAACTTGTAAAGTACCTTAAGAGTCTGCATG 473
           ||||| || || |||||||||   |   |||
Db   33470 GAACTAGTGAAATACCTTAAGGTACCAAATG 33440


RESULT 12
AK074081
LOCUS       AK074081               5564 bp    mRNA    linear   PRI 15-FEB-2002
DEFINITION  Homo sapiens mRNA for FLJ00152 protein.
ACCESSION   AK074081
VERSION     AK074081.1  GI:18676509
KEYWORDS    fis (full insert sequence).
SOURCE      Homo sapiens adult spleen cDNA to mRNA, clone:FLJ00152.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE     The nucleotide sequence of a long cDNA clone isolated from human
            spleen
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 5564)
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research
            Institute, Department of Human Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert and 5'- &
            3'-end one pass sequencing: Research Association for Biotechnology;
            cDNA library construction and clone selection: Kazusa DNA Research
            Institute.
FEATURES             Location/Qualifiers
     source          1. .5564
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="FLJ00152"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /note="vector:pBluescriptII SK plus"
     gene            1. .5564
                     /gene="FLJ00152"
     CDS             <1. .4122
                     /gene="FLJ00152"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /evidence=not_experimental
                     /product="FLJ00152 protein"
                     /protein_id="BAB84907.1"
                     /db_xref="GI:18676510"
                     /translation="GEDASNAMPVIFGKSSGPEFLQEVYTAVTYHNKSPDFYEEVKIK
                     LPAKLTVNHHLLFTFYHISCQQKQGASVETLLGYSWLPILLNERLQTGSYCLPVALEK
                     LPPNYSMHSAEKVPLQNPPIKWAEGHKGVFNIEVQAVSSVHTQDNHLEKFFTLCHSLE
                     SQVTFPIRVLDQKISEMALEHELKLSIICLNSSRLEPLVLFLHLVLDKLFQLSVQPMV
                     IAGQTANFSQFAFESVVAIANSLHNSKDLSKDQHGRNCLLASYVHYVFRLPEVQRDVP
                     KSGAPTALLDPRSYHTYGRTSAAAVSSKLLQARVMSSSNPDLAGTHSAADEEVKNIMS
                     SKIADRNCSRMSYYCSGSSDAPSSPAAPRPASKKHFHEELALQMVVSTGMVRETVFKY
                     AWFFFELLVKSMAQHVHNMDKRDSFRRTRFSDRFMDDITTIVNVVTSEIAALLVKPQK
                     ENEQAEKMNISLAFFLYDLLSLMDRGFVFNLIRHYCSQLSAKLSNLPTLISMRLEFLR
                     ILCSHEHYLNLNLFFMNADTAPTSPCPSISSQNSSSCSSFQDQKIASMFDLTSEYRQQ
                     HFLTGLLFTELAAALDAEGEGISKVQRKAVSAIHSLLSSHDLDPRCVKPEVKVKIAAL
                     YLPLVGIILDALPQLCDFTVADTRRYRTSGSDEEQEGAGAINQNVALAIAGNNFNLKT
                     SGIVLSSLPYKQYNMLNADTTRNLMICFLWIMKNADQSLIRKWIADLPSTQLNRILDL
                     LFICVLCFEYKGKQSSDKVSTQVLQKSRDVKARLEEALLRGEGARGEMMRRRAPGNDR
                     FPGLNENLRWKKEQTHWRQANEKLDKTKAELDQEALISGNLATEAHLIILDMQENIIQ
                     ASSALDCKDSLLGGVLRVLVNSLNCDQSTTYLTHCFATLRALIAKFGDLLFEEEVEQC
                     FDLCHQVLHHCSSSMDVTRSQACATLYLLMRFSFGATSNFARVKMQVTMSLASLVGRA
                     PDFNEEHLRRSLRTILAYSEEDTAMQMTPFPTQVEELLCNLNSILYDTVKMREFQEDP
                     EMLMDLMYRIAKSYQASPDLRLTWLQNMAEKHTKKKCYTEAAMCLVHAAALVAEYLSM
                     LEDHSYLPVGSVSFQNISSNVLEESVVSEDTLSPDEDGVCAGQYFTESGLVGLLEQAA
                     ELFSTGGLYETVNEFYKLFIPILEAHREFRKLTLTHSKLQRAFDSIVNKDHKRMFGTY
                     FRVGFFGSKFGDLDEQEFVYKEPAITKLPEISHRLEAFYGQCFGAEFVEVIKDSTPVD
                     KTKLDPNKAYIQITFVEPYFDEYEMKDRVTYFEKNFNLRRFMYTTPFTLEGRPRGELH
                     EQYRRNTVLTTMHAFPYIKTRISVIQKEEFVLTPIEVAIEDMKKKNPAVSSCH"
BASE COUNT     1533 a   1332 c   1314 g   1385 t
ORIGIN

  Query Match             14.3%;  Score 71.4;  DB 9;  Length 5564;
  Best Local Similarity   51.1%;  Pred. No. 5.3e-11;
  Matches  208;  Conservative    0;  Mismatches  181;  Indels   18;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
             |||  | |||||  |  |    || || |||||||| ||  | |||||  || | ||| ||
Db      82 GTTACATACCATAATAAGTCTCCTGACTTTTATGAAGAAGTGAAAATTAAGCTCCCCGCT 141

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
             ||||    | ||| ||||||||  || |||| |||| |||| |||||||| |  |
Db     142 AAGCTCACAGTAAATCACCACCTCCTGTTCACCTTCTACCATATCAGCTGTCAGCA---- 197

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
                       |||| |  | |    ||| ||||| |   | || || || |||||
Db     198 --------------GAAGCAAGGAGCCTCCGTGGAAACTCTCCTGGGATATTCATGGCTG 243
```

```
Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | ||   | | | | |   |    ||  |     |    |||| ||  |   | |
Db     244 CCAATTCTCTTAAATGAACGTCTTCAAACTGGATCCTACTGTCTCCCAGTTGCCTTGGAA 303

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
               |   | |      |   ||      || ||  | |       |   |||
Db     304 AAATTGCCACCCAACTACTCCATGCATTCTGCTGAGAAAGTCCCATTACAGAATCCTCCC 363

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           ||||| |||| || |||   |||  | |  | || |||     ||    |||||| |
Db     364 ATTAAGTGGGCTGAAGGACATAAGGGAGTATTTAATATTGAAGTGCAAGCTGTTTCTTCT 423

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCA 408
           ||  | || |||||  | ||  |  | || || |||     ||| ||
Db     424 GTACACACCCAGGACAACCACCTGGAGAAGTTCTTCACCCTCTGCCA 470
```

RESULT 13
AK090429
LOCUS AK090429 6422 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens mRNA for FLJ00346 protein.
ACCESSION AK090429
VERSION AK090429.1 GI:21748545
KEYWORDS fis (full insert sequence).
SOURCE Homo sapiens adult spleen cDNA to mRNA, clone:FLJ00346.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jikuya,H., Takano,J., Kikuno,R., Nagase,T. and Ohara,O.
TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 6422)
AUTHORS Jikuya,H., Takano,J., Kikuno,R., Nagase,T. and Ohara,O.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takahiro Nagase, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert and 5'- & 3'-end one pass sequencing: Research Association for Biotechnology; cDNA library construction and clone selection: Kazusa DNA Research Institute.
FEATURES Location/Qualifiers
source 1. .6422
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLJ00346"
/tissue_type="spleen"
/dev_stage="adult"
/note="vector:pBluescriptII SK plus"
gene 1. .6422
/gene="FLJ00346"
CDS <1. .5401
/gene="FLJ00346"
/note="Start codon is not identified."
/codon_start=2
/evidence=not_experimental
/product="FLJ00346 protein"
/protein_id="BAC03410.1"
/db_xref="GI:21748546"
/translation="ENFHCDLNSDQFKGFLRAHTPSVAASSQARSAVFSVTYPSSDIY
LVVKIEKVLQQGEIGDCAEPYTVIKESDGGKSKEKIEKLKLQAESFCQRLGKYRMPFA
WAPISLSSFFNVSTLEREVTDVDSVVGRSSVGERRTLAQSRRLSERALSLEENGVGSN
FKTSTLSVSSFFKQEGDRLSDEDLFKFLADYKRSSSLQRRVKSIPGLLRLEISTAPEI
INCCLTPEMLPVKPFPENRTRPHKEILEFPTREVYVPHTVYRNLLYVYPQRLNFVNKL
ASARNITIKIQFMCGEDASNAMPVIFGKSSGPEFLQEVYTAVTYHNKSPDFYEEVKIK
LPAKLTVNHHLLFTFYHISCQQKQGASVETLLGYSWLPILLNERLQTGSYCLPVALEK
LPPNYSMHSAEKVPLQNPPIKWAEGHKGVFNIEVQAVSSVHTQDNHLEKFFTLCHSLE
SQVTFPIRVLDQKISEMALEHELKLSIICLNSSRLEPLVLFLHLVLDKLFQLSVQPMV
IAGQTANFSQFAFESVVAIANSLHNSKDLSKDQHGRNCLLASYVHYVFRLPEVQRDVP
KSGAPTALLDPRSYHTYGRTSAAAVSSKLLQARVMSSSNPDLAGTHSAADEEVKNIMS
SKIADRNCSRMSYYCSGSSDAPSSPAAPRPASKKHFHEELALQMVVSTGMVRETVFKY
AWFFFELLVKSMAQHVHNMDKRDSFRRTRFSDRFMDDITTIVNVVTSEIAALLVKPQK
ENEQAEKMNISLAFFLYDLLSLMDRGFVFNLIRHYCSQLSAKLSNLPTLISMRLEFLR
ILCSHEHYLNLNLFFMNADTAPTSPCPSISSQNSSSCSSFQDQKIASMFDLTSEYRQQ
HFLTGLLFTELAAALDAEGEGISKVQRKAVSAIHSLLSSHDLDPRCVKPEVKVKIAAL
YLPLVGIILDALPQLCDFTVADTRRYRTSGSDEEQEGAGAINQNVALAIAGNNFNLKT
SGIVLSSLPYKQYNMLNADTTRNLMICFLWIMKNADQSLIRKWIADLPSTQLNRILDL
LFICVLCFEYKGKQSSDKVSTQVLQKSRDVKARLEEALLRGEGARGEMMRRRAPGNDR
FPGLNENLRWKKEQTHWRQANEKLDKTKAELDQEALISGNLATEAHLIILDMQENIIQ
ASSALDCKDSLLGGVLRVLVNSLNCDQSTTYLTHCFATLRALIAKFGDLLFEEEVEQC
FDLCHQVLHHCSSSMDVTRSQACATLYLLMRFSFGATSNFARVKMQVTMSLASLVGRA
PDFNEEHLRRSLRTILAYSEEDTAMQMTPFPTQVEELLCNLNSILYDTVKMREFQEDP
EMLMDLMYRIAKSYQASPDLRLTWLQNMAEKHTKKKCYTEAAMCLVHAAALVAEYLSM
LEDHSYLPVGSVSFQNISSNVLEESVVSEDTLSPDEDGVCAGQYFTESGLVGLLEQAA
ELFSTGGLYETVNEVYKLVIPILEAHREFRKLTLTHSKLQRAFDSIVNKDHKRMFGTY
FRVGFFGSKFGDLDEQEFVYKEPAITKLPEISHRLEAFYGQCFGAEFVEVIKDSTPVD
KTKLDPNKAYIQITFVEPYFDEYEMKDRVTYFEKNFNLRRFMYTTPFTLEGRPRGELH
EQYRRNTVLTTMHAFPYIKTRISVIQKEEFVLTPIEVAIEDMKKKTLQLAVAINQEPP
DAKMLQMVLQGSVGATVNQGPLEVAQVFLAEIPADPKLYRHHNKLRLCFKEFIMRCGE
AVEKNKRLITADQREYQQELKKNYNKLKENLRPMIERKIPELYKPIFRVESQKRDSFH
RSSFRKCETQLSQGS"
BASE COUNT 1785 a 1536 c 1510 g 1591 t
ORIGIN

Query Match 14.3%; Score 71.4; DB 9; Length 6422;
Best Local Similarity 51.1%; Pred. No. 5.5e-11;
Matches 208; Conservative 0; Mismatches 181; Indels 18; Gaps 1;

```
Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
             |||  | |||||  |  |    || || |||||||| ||  | |||||  || | ||| ||
Db     953 GTTACATACCATAATAAGTCTCCTGACTTTTATGAAGAAGTGAAAATTAAGCTCCCCGCT 1012

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
              ||||    | ||| ||||||||  || |||| |||| |||| |||||||| |  |
Db    1013 AAGCTCACAGTAAATCACCACCTCCTGTTCACCTTCTACCATATCAGCTGTCAGCA---- 1068

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
                       |||| |  | |    ||| ||||| |   | || || || |||||
Db    1069 -------------GAAGCAAGGAGCCTCCGTGGAAACTCTCCTGGGATATTCATGGCTG 1114

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | ||   | | | | |   |    ||  |     |    |||| ||  |   | |
Db    1115 CCAATTCTCTTAAATGAACGTCTTCAAACTGGATCCTACTGTCTCCCAGTTGCCTTGGAA 1174

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
               |   | |      |   ||      || ||  | |       |   |||
Db    1175 AAATTGCCACCCAACTACTCCATGCATTCTGCTGAGAAAGTCCCATTACAGAATCCTCCC 1234

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           ||||| |||| || |||   |||  | |  | || |||     ||    |||||| |
Db    1235 ATTAAGTGGGCTGAAGGACATAAGGGAGTATTTAATATTGAAGTGCAAGCTGTTTCTTCT 1294

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCA 408
           ||  | || |||||  | ||  |  | || || |||     ||| ||
Db    1295 GTACACACCCAGGACAACCACCTGGAGAAGTTCTTCACCCTCTGCCA 1341
```

RESULT 14
AX174569
LOCUS AX174569 7215 bp DNA linear PAT 04-JUL-2001
DEFINITION Sequence 1 from Patent WO0142296.
ACCESSION AX174569
VERSION AX174569.1 GI:14598200
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7215)
AUTHORS Lu,P., Garman,J.D. and Candia,A.F.

```
  TITLE     Clasp-5 transmembrane protein
  JOURNAL   Patent: WO 0142296-A 1 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7215
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="full length human CLASP-5 cDNA"
     CDS             112. .6159
                     /note="human CLASP-5"
                     /codon_start=1
                     /protein_id="CAC43733.1"
                     /db_xref="GI:14598201"
                     /translation="MTHLNSLDVQLAQELGDFTDDDLDVVFTPKECRTLQPSLPEEGV
                     ELDPHVRDCVQTYIREWLIVNRKNQGSPEICGFKKTGSRKDFHKTLPKQTFESETLEC
                     SEPAAQAGPRHLNVLCDVSGKGPVTACDFDLRSLQPDKRLENLLQQVSAEDFEKQNEE
                     ARRTNRQAELFALYPSVDEEDAVEIRPVPECPKEHLGNRILVKLLTLKFEIEIEPLFA
                     SIALYDVKERKKISENFHCDLNSDQFKGFLRAHTPSVAASSQARSAVFSVTYPSSDIY
                     LVVKIEKVLQQGDIGDCAEPYTVIKESDGGKSKEKIEKLKLQAESFCQRLGKYRMPFA
                     WAPISLSSFFNVSTLEREVTDVDSVVGRSPVGERRTLAQSRRLSERALSLEENGVGSN
                     FKTSTLSVSSFFKQEGDRLSDEDLFKFLADYKRSSSLQRRVKSIPGLLRLEISTAPEI
                     INCCLTPEMLPVKPFPENRTRPHKEILEFPTREVYVPHTVYRNLLYVYPQRLNFVNKL
                     ASARNITIKIQFMCGEDASNAMPVIFGKSSGPEFLQEVYTAVTYHNKSPDFYEEVKIK
                     LPAKLTVNHHLLFTFYHISCQQKQGASVETLLGYSWLPILLNERLQTGSYCLPVALEK
                     LPPNYSMHSAEKVPLQNPPIKWAEGHKGVFNIEVQAVSSVHTQDNHLEKFFTLCHSLE
                     SQVTFPIRVLDQKISEMALEHELKLSIICLNSSRLEPLVLFLHLVLDKLFQLSVQPMV
                     IAGQTANFSQFAFESVVAIANSLHNSKDLSKDQHGRNCLLASYVHYVFRLPEVQRDVP
                     KSGAPTALLDPRSYHTYGRTSAAAVSSKLLQARVMSSSNPDLAGTHSAADEEVKNIMS
                     SKIADRNCSRMSYYCSGSSDAPSSPAAPRPASKKHFHEELALQMVVSTGMVKSMAQHV
                     HNMDKRDSFRRTRFSDRFMDDITTIVNVVTSEIAALLVKPQKENEQAEKMNISLAFFL
                     YDLLSLMDRGFVFNLIRHYCSQLSAKLSNLPTLISMRLEFLRILCSHEHYLNLNLFFM
                     NADTAPTSPCPSISSQNSSSCSSFQDQKIASMFDLTSEYRQQHFLTGLLFTELAAALD
                     AEGEGISKVQRKAVSAIHSLLSSHDLDPRCVKPEVKVKIAALYLPLVGIILDALPQLC
                     DFTVADTRRYRTSGSDEEQEGAGAINQNVALAIAGNNFNLKTSGIVLSSLPYKQYNML
                     NADTTRNLMICFLWIMKNADQSLIRKWIADLPSTQLNRILDLLFICVLCFEYKGKQSS
                     DKVSTQVLQKSRDVKARLEEALLRGEGARGEMMRRRAPGNDRFPGLNENLRWKKEQTH
                     WRQANEKLDKTKAELDQEALISGNLATEAHLIILDMQENIIQASSALDCKDSLLGGVL
                     RVLVNSLNCDQSTTYLTHCFATLRALIAKFGDLLFEEEVEQCFDLCHQVLHHCSSSMD
                     VTRSQACATLYLLMRFSFGATSNFARVKMQVTMSLASLVGRAPDFNEEHLRRSLRTIL
                     AYSEEDTAMQMTPFPTQVEELLCNLNSILYDTVKMREFQEDPEMLMDLMYRIAKSYQA
                     SPDLRLTWLQNMAEKHTKKKCYTEAAMCLVHAAALVAEYLSMLEDHSYLPVGSVSFQN
                     ISSNVLEESVVSEDTLSPDEDGVCAGQYFTESGLVGLLEQAAELFSTGGLYETVNEVY
                     KLVIPILEAHREFRKLTLTHSKLQRAFDSIVNKDHKRMFGTYFRVGFFGSKFGDLDEQ
                     EFVYKEPAITKLPEISHRLEAFYGQCFGAEFVEVIKDSTPVDKTKLDPNKAYIQITFV
                     EPYFDEYEMKDRVTYFEKNFNLRRFMYTTPFTLEGRPRGELHEQYRRNTVLTTMHAFP
                     YIKTRISVIQKEEFVLTPIEVAIEDMKKKTLQLAVAINQEPPDAKMLQMVLQGSVGAT
                     VNQGPLEVAQVFLAEIPADPKLYRHHNKLRLCFKEFIMRCGEAVEKNKRLITADQREY
                     QQELKKNYNKLKENLRPMIERKIPELYKPIFRVESQKRDSFHRSSFRKCETQLSQGS"
BASE COUNT     2005 a   1734 c   1721 g   1755 t
ORIGIN

  Query Match             14.3%;  Score 71.4;  DB 6;  Length 7215;
  Best Local Similarity   51.1%;  Pred. No. 5.6e-11;
  Matches  208;  Conservative    0;  Mismatches  181;  Indels   18;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           |||  | |||||  |  |    || || ||||||||  ||  | |||||  || | ||| ||
Db    1759 GTTACATACCATAATAAGTCTCCTGACTTTTATGAAGAAGTGAAAATTAAGCTCCCCGCT 1818

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
            ||||    | ||| ||||||||  || |||| |||| |||| ||||||||  |  |
Db    1819 AAGCTCACAGTAAATCACCACCTCCTGTTCACCTTCTACCATATCAGCTGTCAGCA---- 1874

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
                      |||| |  | |    ||| ||||| |   | || || || |||||
Db    1875 --------------GAAGCAAGGAGCCTCCGTGGAAACTCTCCTGGGATATTCATGGCTG 1920

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | ||   | | | | |  |    ||  |     |     |||| ||  |   | |
Db    1921 CCAATTCTCTTAAATGAACGTCTTCAAACTGGATCCTACTGTCTCCCAGTTGCCTTGGAA 1980

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
                 |   | |      |   ||      || ||  | |       |   |||
Db    1981 AAATTGCCACCCAACTACTCCATGCATTCTGCTGAGAAAGTCCCATTACAGAATCCTCCC 2040

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           |||||  ||||  || |||   |||  | |  | || |||     ||   ||||||  |
Db    2041 ATTAAGTGGGCTGAAGGACATAAGGGAGTATTTAATATTGAAGTGCAAGCTGTTTCTTCT 2100

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCA 408
           ||  | || |||||  | ||  |  | || || |||     ||| ||
Db    2101 GTACACACCCAGGACAACCACCTGGAGAAGTTCTTCACCCTCTGCCA 2147


RESULT 15
HSM803577
LOCUS       HSM803577               9298 bp .   mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp667O1117 (from clone DKFZp667O1117).
ACCESSION   AL832270
VERSION     AL832270.1  GI:21732817
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9298)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp667O1117) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .9298
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp667O1117"
                     /tissue_type="lymph node"
                     /clone_lib="667 (synonym: hlno2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_signal    9259. .9264
     polyA_site      9281
BASE COUNT     2546 a   2239 c   2161 g   2352 t
ORIGIN

  Query Match             14.3%;  Score 71.4;  DB 9;  Length 9298;
  Best Local Similarity   51.1%;  Pred. No. 5.9e-11;
  Matches  208;  Conservative    0;  Mismatches  181;  Indels   18;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           |||  | |||||  |  |    || || ||||||||  ||  | |||||  || | ||| ||
Db    3805 GTTACATACCATAATAAGTCTCCTGACTTTTATGAAGAAGTGAAAATTAAGCTCCCCGCT 3864

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
            ||||    | ||| ||||||||  || |||| |||| |||| ||||||||  |  |
Db    3865 AAGCTCACAGTAAATCACCACCTCCTGTTCACCTTCTACCATATCAGCTGTCAGCA---- 3920

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
                      |||| |  | |    ||| ||||| |   | || || || |||||
Db    3921 --------------GAAGCAAGGAGCCTCCGTGGAAACTCTCCTGGGATATTCATGGCTG 3966

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | ||   | | | | |  |    ||  |     |     |||| ||  |   | |
Db    3967 CCAATTCTCTTAAATGAACGTCTTCAAACTGGATCCTACTGTCTCCCAGTTGCCTTGGAA 4026

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
                 |   | |      |   ||      || ||  | |       |   |||
Db    4027 AAATTGCCACCCAACTACTCCATGCATTCTGCTGAGAAAGTCCCATTACAGAATCCTCCC 4086
```

```
Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           ||||| ||||  || |||   |||  | |  | || |||      ||    |||||| |
Db    4087 ATTAAGTGGGCTGAAGGACATAAGGGAGTATTTAATATTGAAGTGCAAGCTGTTTCTTCT 4146

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCA 408
           ||  | || |||||  | ||  |  | || || |||     ||| ||
Db    4147 GTACACACCCAGGACAACCACCTGGAGAAGTTCTTCACCCTCTGCCA 4193


Search completed: February  7, 2003, 08:03:14
Job time : 1332.67 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:36:04 ; Search time 127.83 Seconds
(without alignments)
8808.598 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttttacaccatcaccaaa.....aggccacgtgatgatcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	500	100.0	4807	21	AAC87972	Human CLASP-2 nucl
2	500	100.0	4807	21	AAC87973	Human CLASP-2A nuc
3	500	100.0	4807	24	ABK84966	DNA encoding cadhe
4	500	100.0	4807	24	ABK84973	DNA encoding cadhe
5	499	99.8	4806	24	ABK84964	DNA encoding cadhe
6	499	99.8	5048	24	ABK84965	DNA encoding cadhe
7	498.4	99.7	5862	24	ABK84970	DNA encoding cadhe
8	498.4	99.7	6816	21	AAC74524	Human ORFX ORF79 p
9	498.4	99.7	7506	22	AAD19118	Angiogenesis assoc
10	496.8	99.4	6791	24	ABK85003	DNA encoding cadhe
11	486.4	97.3	4898	21	AAC87974	Preliminary CLASP-
12	486.4	97.3	4898	21	AAC87975	Preliminary CLASP-
13	486.4	97.3	4898	21	AAC87976	Preliminary CLASP-
14	486.4	97.3	4898	21	AAC87977	Preliminary CLASP-
15	486.4	97.3	4898	21	AAC87978	Preliminary CLASP-
16	486.4	97.3	4898	21	AAC87979	Preliminary CLASP-
17	486.4	97.3	4898	21	AAC87980	Preliminary CLASP-
18	486.4	97.3	4898	21	AAC87981	Preliminary CLASP-
19	486.4	97.3	4898	24	ABK84992	DNA encoding cadhe
20	486.4	97.3	4898	24	ABK84993	DNA encoding cadhe
21	486.4	97.3	4898	24	ABK84994	DNA encoding cadhe
22	486.4	97.3	4898	24	ABK84995	DNA encoding cadhe
23	486.4	97.3	4898	24	ABK84996	DNA encoding cadhe
24	486.4	97.3	4898	24	ABK84997	DNA encoding cadhe
25	486.4	97.3	4898	24	ABK84998	DNA encoding cadhe
26	486.4	97.3	4898	24	ABK84999	DNA encoding cadhe
27	204.4	40.9	6454	22	AAS08334	Human cDNA encodin
28	173.8	34.8	5688	21	AAC87969	Human CLASP relate
29	173.8	34.8	5688	21	AAA14825	DNA encoding a hum
30	173.8	34.8	7277	24	AAS18951	Human cDNA encodin
31	117.2	23.4	5214	21	AAC87968	Mouse CLASP-1 nucl
32	117.2	23.4	5214	21	AAA14824	DNA encoding a mur
33	85	17.0	4393	22	AAS08357	Human cDNA encodin
34	71.4	14.3	7215	22	AAS07373	Human cDNA encodin
c 35	69.8	14.0	665	24	ABK53548	Human eosinophil-m
36	54.2	10.8	5955	23	ABL11149	Drosophila melanog
37	46	9.2	6372	22	AAH43851	Human CLASP-7 enco
38	42.4	8.5	5589	23	ABL17737	Drosophila melanog
c 39	42.4	8.5	9389	23	ABL17736	Drosophila melanog
40	41	8.2	6828	22	AAH41934	Human CLASP-3 cDNA
41	40.4	8.1	211	22	AAH43868	Human CLASP-7 intr
42	40.4	8.1	3023	22	AAH17551	Human cDNA sequenc
c 43	38.6	7.7	8942	23	ABL11148	Drosophila melanog
44	34.6	6.9	6691	20	AAX13025	Enterococcus faeca
c 45	34	6.8	1851	21	AAA52623	Eosinophil activat

ALIGNMENTS

RESULT 1
AAC87972
ID AAC87972 standard; cDNA; 4807 BP.
XX
AC AAC87972;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2 nucleotide sequence.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.

PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PSDB; AAB36527.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Example 1; Fig 1; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence encodes human CLASP-2, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 500; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480

Qy     481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db     481 AGGCCACGTGATGATCGCCT 500
```

RESULT 2
AAC87973
ID AAC87973 standard; cDNA; 4807 BP.
XX
AC AAC87973;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2A nucleotide sequence.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PSDB; AAB36528.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Example 1; Fig 2B; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat

CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence encodes human CLASP-2A, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 500; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180

Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300

Qy    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360

Qy    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420

Qy    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480

Qy    481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    481 AGGCCACGTGATGATCGCCT 500
```

RESULT 3
ABK84966
ID ABK84966 standard; cDNA; 4807 BP.
XX
AC ABK84966;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #1.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61672.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 500; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180

Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300

Qy    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360

Qy    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420

Qy    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480

Qy    481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    481 AGGCCACGTGATGATCGCCT 500
```

RESULT 4
ABK84973
ID ABK84973 standard; DNA; 4807 BP.
XX
AC ABK84973;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61686.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Example 4; Figure 6A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences
CC and PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 500; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.1e-157;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180

Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480

Qy     481 AGGCCACGTGATGATCGCCT 500
                ||||||||||||||||||||
Db     481 AGGCCACGTGATGATCGCCT 500
```

RESULT 5
ABK84964
ID ABK84964 standard; cDNA; 4806 BP.
XX
AC ABK84964;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP).
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61670.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 1; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other;

```
  Query Match             99.8%;  Score 499;  DB 24;  Length 4806;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-157;
  Matches  499;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 60

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 120

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 180

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 240

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 300

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 360

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 420

Qy     422 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 480

Qy     482 GGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
```

```
Db    481 GGCCACGTGATGATCGCCT 499


RESULT 6
ABK84965
ID   ABK84965 standard; cDNA; 5048 BP.
XX
AC   ABK84965;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200231117-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
PA   (ARBO-) ARBOR VITA CORP.
PA   (GARM/) GARMAN J D.
PA   (CAND/) CANDIA A F.
XX
PI   Lu PS;
XX
DR   WPI; 2002-416861/44.
DR   P-PSDB; ABG61671.
XX
PT   New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT   an immune response, and for treating multiple sclerosis, rheumatoid
PT   arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT   and sepsis -
XX
PS   Disclosure; Figure 2; 245pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising an amino
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g., agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. CLASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC   PCR primers of the invention.
XX
SQ   Sequence 5048 BP; 1403 A; 1187 C; 1215 G; 1243 T; 0 other;

  Query Match             99.8%;  Score 499;  DB 24;  Length 5048;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-157;
  Matches  499;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 60

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 120

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 180

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 240

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 300

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 360

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 420

Qy     422 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 480

Qy     482 GGCCACGTGATGATCGCCT 500
           ||||||||||||||||||
Db     481 GGCCACGTGATGATCGCCT 499


RESULT 7
ABK84970
ID   ABK84970 standard; cDNA; 5862 BP.
XX
AC   ABK84970;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) isoform #5.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
```

KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61676.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 5862 BP; 1670 A; 1327 C; 1341 G; 1524 T; 0 other;

Query Match 99.7%; Score 498.4; DB 24; Length 5862;
Best Local Similarity 99.8%; Pred. No. 1.2e-156;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 554

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 614

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 674

Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 734

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    735 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 794

Qy    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 854

Qy    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 914

Qy    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 974

Qy    481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    975 AGGCCACGTGATGATCGCCT 994
```

RESULT 8
AAC74524
ID AAC74524 standard; cDNA; 6816 BP.
XX
AC AAC74524;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF79 polynucleotide sequence SEQ ID NO:157.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX

```
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB40315.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 502-506; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 6816 BP; 1977 A; 1523 C; 1557 G; 1757 T; 2 other;

  Query Match             99.7%;  Score 498.4;  DB 21;  Length 6816;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-156;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 1502

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 1562

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 1622

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 1682

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    1683 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 1742

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 1802

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 1862

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 1922

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db    1923 AGGCCACGTGATGATCGCCT 1942


RESULT 9
AAD19118
ID   AAD19118 standard; cDNA; 7506 BP.
XX
AC   AAD19118;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Angiogenesis associated human thyroid regulated transcript (hTRG) cDNA.
XX
KW   Angiogenesis associated protein; AAP; cytostatic; cardiant; gene therapy;
KW   ophthalmological; vulnerary; myocardial infarction; macular degeneration;
KW   diabetic retinopathy; angiogenesis; wound healing; prophylactic; vaccine;
KW   rheumatoid arthritis; psoriasis; drug screening; tumour; transplantation;
KW   cancer; therapeutic; diagnostic; human; thyroid regulated transcript;
KW   TRG; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..6366
FT                   /*tag=  a
FT                   /product= "human thyroid regulated transcript (hTRG)"
XX
PN   WO200170808-A2.
XX
PD   27-SEP-2001.
XX
PF   22-MAR-2001; 2001WO-US09609.
XX
PR   22-MAR-2000; 2000US-191134P.
XX
PA   (CURA-) CURAGEN CORP.
PA   (GETH ) GENENTECH INC.
XX
PI   Rastelli LK,  Gerritsen M;
XX
DR   WPI; 2001-602775/68.
DR   P-PSDB; AAE11889.
XX
PT   Novel angiogenesis associated polypeptides and polynucleotides encoding
PT   the polypeptides, useful for modulating angiogenesis and for treating
PT   tumors and cancers -
XX
PS   Claim 6; Page 18-21; 159pp; English.
XX
CC   The invention relates to angiogenesis associated proteins (AAP) and their
CC   corresponding cDNA molecules, which are useful for modulating
CC   angiogenesis. AAP proteins and nucleic acids are useful for promoting
CC   wound healing, for example after organ transplantation, and in the
CC   treatment of tumours, myocardial infarction, cancers, diabetic
CC   retinopathy, macular degeneration, psoriasis and rheumatoid arthritis.
CC   AAP proteins and DNA's are useful in potential prophylactic and
CC   therapeutic applications implicated in a variety of disorders including
CC   those related to angiogenesis, and also in diagnostic applications.
CC   AAP cDNA is also useful in gene therapy. The invention also relates to
CC   a method for screening a tissue sample for tumourigenic potential. AAP
CC   proteins are used to screen drugs or compounds that modulate AAP activity
CC   or expression as well as treating disorders characterised by insufficient
CC   or excessive production of AAP or production of AAP forms that have
CC   decreased or aberrant activity compared to the wild type protein, or
```

```
CC   modulate biological function that involve AAP. The present cDNA sequence
CC   encodes human thyroid regulated transcript (hTRG) protein which is an
CC   angiogenesis associated protein (AAP) of the invention. Human TRG is
CC   upregulated in the in vitro model of angiogenesis and is likely to be
CC   involved in signal transduction between receptors and kinases. Modulation
CC   of hTRG is useful to treat diseases related to thyroid stimulating
CC   hormone (TSH) imbalance.
XX
SQ   Sequence 7506 BP; 2147 A; 1689 C; 1764 G; 1906 T; 0 other;

  Query Match           99.7%;  Score 498.4;  DB 22;  Length 7506;
  Best Local Similarity  99.8%;  Pred. No. 1.4e-156;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2142 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 2201

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2202 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 2261

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2262 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 2321

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2322 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 2381

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2382 CCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 2441

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2442 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 2501

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2502 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 2561

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2562 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 2621

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db    2622 AGGCCACGTGATGATCGCCT 2641


RESULT 10
ABK85003
ID   ABK85003 standard; DNA; 6791 BP.
XX
AC   ABK85003;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) from exon 2.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200231117-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
PA   (ARBO-) ARBOR VITA CORP.
PA   (GARM/) GARMAN J D.
PA   (CAND/) CANDIA A F.
XX
PI   Lu PS;
XX
DR   WPI; 2002-416861/44.
DR   P-PSDB; ABG61702.
XX
PT   New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT   an immune response, and for treating multiple sclerosis, rheumatoid
PT   arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT   and sepsis -
XX
PS   Example 1; Figure 11; 245pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising an amino
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g., agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. CLASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences
CC   and PCR primers of the invention.
XX
SQ   Sequence 6791 BP; 1968 A; 1545 C; 1574 G; 1704 T; 0 other;

  Query Match           99.4%;  Score 496.8;  DB 24;  Length 6791;
  Best Local Similarity  99.6%;  Pred. No. 4.4e-156;
  Matches  498;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1985 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 2044
```

```
Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2045 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 2104

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2105 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 2164

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    2165 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGTA 2224

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    2225 CCTTCCTTCGGGCCATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 2284

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2285 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 2344

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2345 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 2404

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2405 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 2464

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db    2465 AGGCCACGTGATGATCGCCT 2484
```

RESULT 11
AAC87974
ID AAC87974 standard; cDNA; 4898 BP.
XX
AC AAC87974;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #1.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PSDB; AAB36529.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10A; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 97.3%; Score 486.4; DB 21; Length 4898;
Best Local Similarity 99.6%; Pred. No. 1.2e-152;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 152

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 212

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 272

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 332

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 392

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 452

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 AG-GGATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 511

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     512 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 571

Qy     481 AGGCCACGTGATGATCGCCT 500
```

```
                    ||||||||||||||||||||
Db    572 AGGCCACGTGATGATCGCCT 591


RESULT 12
AAC87975
ID   AAC87975 standard; cDNA; 4898 BP.
XX
AC   AAC87975;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #2.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10B; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             97.3%;  Score 486.4;  DB 21;  Length 4898;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-152;
  Matches  498;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 152

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 212

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 272

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 332

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 392

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 452

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 AG-GGATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 511

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     512 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 571

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db     572 AGGCCACGTGATGATCGCCT 591


RESULT 13
AAC87976
ID   AAC87976 standard; cDNA; 4898 BP.
XX
AC   AAC87976;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #3.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
```

```
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10C; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             97.3%;  Score 486.4;  DB 21;  Length 4898;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-152;
  Matches  498;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 152

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 212

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 272

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 332

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 392

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 452

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 AG-GGATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 511

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     512 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 571

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db     572 AGGCCACGTGATGATCGCCT 591


RESULT 14
AAC87977
ID   AAC87977 standard; cDNA; 4898 BP.
XX
AC   AAC87977;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #4.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10D; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
```

CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 97.3%; Score 486.4; DB 21; Length 4898;
Best Local Similarity 99.6%; Pred. No. 1.2e-152;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 152

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 212

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 272

Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 332

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 392

Qy    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 452

Qy    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 AG-GGATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 511

Qy    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    512 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 571

Qy    481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    572 AGGCCACGTGATGATCGCCT 591
```

RESULT 15
AAC87978
ID AAC87978 standard; cDNA; 4898 BP.
XX
AC AAC87978;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #5.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10E; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 97.3%; Score 486.4; DB 21; Length 4898;
Best Local Similarity 99.6%; Pred. No. 1.2e-152;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy      1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 152

Qy     61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 212

Qy    121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 272
```

```
Qy    181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 332

Qy    241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 392

Qy    301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 452

Qy    361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
          || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 AG-GGATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 511

Qy    421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    512 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 571

Qy    481 AGGCCACGTGATGATCGCCT 500
          ||||||||||||||||||||
Db    572 AGGCCACGTGATGATCGCCT 591
```

Search completed: February 7, 2003, 07:07:37
Job time : 137.83 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 ; Search time 28.6285 Seconds
(without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttttacaccatcaccaaa.....aggccacgtgatgatcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	30.4	6.1	1005	4	US-08-936-165A-228	Sequence 228, App
c	2	30.4	6.1	1920	4	US-09-160-496-4	Sequence 4, Appli
	3	30.2	6.0	1972	3	US-08-961-083-203	Sequence 203, App
	4	30.2	6.0	14231	4	US-08-961-527-81	Sequence 81, Appl
c	5	30	6.0	2520	1	US-08-405-254-9	Sequence 9, Appli
	6	29.4	5.9	442	4	US-09-071-035-443	Sequence 443, App
	7	29.4	5.9	606	4	US-09-071-035-441	Sequence 441, App
c	8	29.4	5.9	1006	3	US-08-924-747-15	Sequence 15, Appl
c	9	29.4	5.9	1006	4	US-09-247-373B-15	Sequence 15, Appl
c	10	29.4	5.9	1006	4	US-09-296-715-15	Sequence 15, Appl
c	11	29.2	5.8	726	3	US-08-660-645A-11	Sequence 11, Appl
c	12	29.2	5.8	726	3	US-09-298-718-11	Sequence 11, Appl
c	13	29.2	5.8	726	4	US-09-546-969-11	Sequence 11, Appl
c	14	29.2	5.8	726	4	US-08-980-832-28	Sequence 28, Appl
c	15	29.2	5.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	16	29.2	5.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	17	29	5.8	856	1	US-08-117-373-10	Sequence 10, Appl
	18	29	5.8	1542	4	US-09-008-271A-13	Sequence 13, Appl
c	19	29	5.8	4084	2	US-08-568-459A-1	Sequence 1, Appli
c	20	29	5.8	4084	2	US-08-487-826B-1	Sequence 1, Appli
c	21	29	5.8	4084	4	US-09-210-288-1	Sequence 1, Appli
c	22	29	5.8	4084	6	5198347-5	Patent No. 5198347
c	23	29	5.8	4156	4	US-08-961-527-211	Sequence 211, App
	24	28.8	5.8	1747	1	US-07-808-455A-2	Sequence 2, Appli
c	25	28.6	5.7	1643	1	US-08-383-750-3	Sequence 3, Appli
c	26	28.6	5.7	1643	3	US-08-352-678-3	Sequence 3, Appli
c	27	28.6	5.7	1643	5	PCT-US93-09636-3	Sequence 3, Appli
	28	28.4	5.7	1072	1	US-07-971-096-1	Sequence 1, Appli
	29	28.4	5.7	1072	1	US-08-175-096-1	Sequence 1, Appli
	30	28.2	5.6	608	4	US-09-385-982-236	Sequence 236, App
c	31	28.2	5.6	18609	4	US-08-943-731-1	Sequence 1, Appli
c	32	28	5.6	488	4	US-09-385-982-368	Sequence 368, App
c	33	28	5.6	111282	4	US-09-754-250-3	Sequence 3, Appli
c	34	27.8	5.6	2970	4	US-09-110-517-3	Sequence 3, Appli
c	35	27.8	5.6	3826	4	US-09-302-620B-90	Sequence 90, Appl
c	36	27.8	5.6	9412	2	US-08-955-138-1	Sequence 1, Appli
	37	27.6	5.5	1463	4	US-09-399-913-1	Sequence 1, Appli
	38	27.6	5.5	1463	4	US-09-298-731-1	Sequence 1, Appli
	39	27.6	5.5	1540	4	US-09-399-913-9	Sequence 9, Appli
	40	27.6	5.5	1540	4	US-09-298-731-9	Sequence 9, Appli
	41	27.6	5.5	1907	4	US-09-399-913-5	Sequence 5, Appli
	42	27.6	5.5	1907	4	US-09-298-731-5	Sequence 5, Appli
c	43	27.6	5.5	2624	1	US-08-032-382B-1	Sequence 1, Appli
c	44	27.4	5.5	518	4	US-09-123-912-87	Sequence 87, Appl
c	45	27.4	5.5	518	4	US-09-643-597-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-08-936-165A-228/c
; Sequence 228, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 228:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-936-165A-228

  Query Match             6.1%;  Score 30.4;  DB 4;  Length 1005;
  Best Local Similarity   62.5%;  Pred. No. 2.2;
  Matches   65;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGA---TGAGATTAAAATAGAGTTGCCC 58
           ||| | ||||| ||  || ||| | ||||| ||||   | ||  |  |||  || | | |
Db     439 GTTGTCCACCAGCAATAATACCTATAATTTCATGAGGGTCAGTATTGAATGTAGGTACAC 380

Qy      59 ACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCA 102
           |||| |||||||  ||  ||||   | || | |   | | | ||
Db     379 ACTCCGCTGCATCTAAGACACCCAATCTTCCACTTGTACCTGCA 336


RESULT 2
US-09-160-496-4/c
; Sequence 4, Application US/09160496
; Patent No. 6346613
; GENERAL INFORMATION:
;  APPLICANT: O'Mahony, Daniel J
;  APPLICANT: Cagney, Gerard
;  TITLE OF INVENTION: Composition and Method for Enhancing Paracellular
;  TITLE OF INVENTION: Transport across Cell Layers
;  FILE REFERENCE: Docket No. 6346613: 98.1070.US
;  CURRENT APPLICATION NUMBER: US/09/160,496
;  CURRENT FILING DATE: 1998-09-24
;  EARLIER APPLICATION NUMBER: US 60/059,644
;  EARLIER FILING DATE: 1997-09-24
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Gallus gallus
;  PUBLICATION INFORMATION:
;   TITLE: Occludin:  A novel integral membrane protein localizing
;   TITLE: at tight junctions
;   JOURNAL: J. Cell Biol.
;   VOLUME: 123
;   ISSUE: 6
;   PAGES: 1777-1788
;   DATE: Dec 1993
;   DATABASE ACCESSION NUMBER: D21837
US-09-160-496-4

  Query Match             6.1%;  Score 30.4;  DB 4;  Length 1920;
  Best Local Similarity   53.3%;  Pred. No. 3.1;
  Matches   64;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||||||||| |  |||| | | |||  ||   ||      | |  | ||  || |  |
Db     842 CCCCTCCTGCACCACGGGCGCGCGGTCCCAGTAGATGTTGGCTTTGCCGTAGCGCCAGAT 783

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           ||| || || |   ||| ||| |  |||    || |   |    | || | | |  ||||
Db     782 CTTACTGCGCGTCTTCTGGGCGAAGAAGCAGATGAGGCAGAGCAGGATGACGATGAGGAA 723


RESULT 3
US-08-961-083-203
; Sequence 203, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  203:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1972 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-083-203

  Query Match             6.0%;  Score 30.2;  DB 3;  Length 1972;
  Best Local Similarity   53.9%;  Pred. No. 3.7;
  Matches   62;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy     238 GAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCC 297
           ||| |  | ||   ||| |||  | |      |   |    ||||  | | ||||||||
Db      44 GAATCGCCATTGTTGCTTTCTCCGTTGTATGGCAATGATAATGGTAACGGATTATGGTGG 103

Qy     298 GGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTG 352
           || ||   | ||    ||    || |||| | | ||||| || |  |  ||| ||
Db     104 GGGAACACATTGAAGGGAGCATGGGAAGCTATTCCTGAAGATGTAAAGCCATATG 158


RESULT 4
US-08-961-527-81
; Sequence 81, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
```

```
;       REFERENCE/DOCKET NUMBER:  PB340P1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 309-8504
;       TELEFAX:  (301) 309-8512
;   INFORMATION FOR SEQ ID NO:  81:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14231 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-961-527-81

  Query Match             6.0%;  Score 30.2;  DB 4;  Length 14231;
  Best Local Similarity   53.9%;  Pred. No. 11;
  Matches   62;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy     238 GAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCC 297
           ||| |  | ||   ||| |||  | |     |   |    ||||  | | ||||||||
Db    9070 GAATCGCCATTGTTGCTTTCTCCGTTGTATGGCAATGATAATGGTAACGGATTATGGTGG 9129

Qy     298 GGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTG 352
           || ||   | ||    ||    || |||| | | ||||| || |  |  ||| ||
Db    9130 GGGAACACATTGAAGGGAGCATGGGAAGCTATTCCTGAAGATGTAAAGCCATATG 9184


RESULT 5
US-08-405-254-9/c
; Sequence 9, Application US/08405254
; Patent No. 5773288
;  GENERAL INFORMATION:
;    APPLICANT:  BRIGGS, STEVEN P.
;    APPLICANT:  BENSON, ROBERT J.
;    TITLE OF INVENTION:  PLANT GENES AFFECTING GIBBERELLIC ACID
;    TITLE OF INVENTION:  BIOSYNTHESIS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  3000 K STREET, NW, SUITE 500
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,254
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,465
;      FILING DATE:  17-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, STEPHEN A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/299/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2520 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-405-254-9

  Query Match             6.0%;  Score 30;  DB 1;  Length 2520;
  Best Local Similarity   47.4%;  Pred. No. 5;
  Matches   90;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     149 GTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTG 208
           || || |   ||   |  ||   || |||     | ||||   |    |  ||  ||  |
Db    2219 GTTGTCGTCGCCGTGGCCGGGGCCTGCTGCTGCGCTCTCCACCACCCGGCGAGCCTGAGG 2160

Qy     209 ACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAA 268
                 |   | | || ||  | || |    | |||| | |    |  ||  |  | |
Db    2159 CTGCCGCGGGGCCCCGTCGCGCCCGCGCCTCCTCCTCCTCCTGCTGCTGCTGCTTTGCCA 2100

Qy     269 GAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCA 328
           | |  ||||| | |  |   |||| || ||      |   | || || |||||||  |
Db    2099 CGGATACGGATGCGGATACGGCATGGACCTGATTGACGAGCTGCAGCTGCAGGCAAGTGA 2040

Qy     329 CTGCTGAAAA 338
           || |||||
Db    2039 AAGCAGAAAA 2030


RESULT 6
US-09-071-035-443
; Sequence 443, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  443:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  442 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-443

  Query Match             5.9%;  Score 29.4;  DB 4;  Length 442;
  Best Local Similarity   48.5%;  Pred. No. 3;
  Matches   81;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy     300 AAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTA 359
           |||  ||  |  | ||||     |    | |   |  | |||||        |||| | |
Db     240 AAAAAAATGGAATTGATGTTAAAACTGGAGTAGCGTTAGTTTCCGTAGGAAAGGTTACAA 299

Qy     360 CAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAAT 419
           |  | ||||  | ||||| |||  | ||  |||  |   | | ||| |  |||  | |
Db     300 CGATTTATAAAAATGATCAACATGAATATGCTTTACTTTTGGATTGTGAACAAAATGGAA 359

Qy     420 CTGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGT 466
```

```
Db    360 CGCAGACACGAGTGTTACTTTTGGCTAAGGTGAAGAACAATAAAATT 406


RESULT 7
US-09-071-035-441
; Sequence 441, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  441:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  606 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-441

  Query Match             5.9%;  Score 29.4;  DB 4;  Length 606;
  Best Local Similarity   48.5%;  Pred. No. 3.6;
  Matches   81;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy     300 AAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTA 359
           |||  ||  |  | ||||     |    | |   |  | |||||        ||||  | |
Db     401 AAAAAAATGGAATTGATGTTAAAACTGGAGTAGCGTTAGTTTCCGTAGGAAAGGTTACAA 460

Qy     360 CAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAAT 419
           |  | ||||   | |||||  |||  | ||  |||  |   | | |||  |  |||  | |
Db     461 CGATTTATAAAAATGATCAACATGAATATGCTTTACTTTTGGATTGTGAACAAAATGGAA 520

Qy     420 CTGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGT 466
           |     | | ||  |||      |     || ||| ||  ||| | |
Db     521 CGCAGACACGAGTGTTACTTTTGGCTAAGGTGAAGAACAATAAAATT 567


RESULT 8
US-08-924-747-15/c
; Sequence 15, Application US/08924747
; Patent No. 6063570
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,747
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1006 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SR1.PK0011.D6
US-08-924-747-15

  Query Match             5.9%;  Score 29.4;  DB 3;  Length 1006;
  Best Local Similarity   60.8%;  Pred. No. 4.8;
  Matches   48;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      63 AGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAA 122
           ||||  ||      | | || |  | |||||  |||| |   |||||||||| |||  |
Db     412 AGCTCAATTGCTTCCTCAACATTCTTCTCACGCTCTTTCTCATCAGCTGTGAAAACGGAT 353

Qy     123 GTAAAGGAAGCACGAAGAA 141
           |  | | ||||| || ||
Db     352 TTCCATGCAGCACCAAAAA 334


RESULT 9
US-09-247-373B-15/c
; Sequence 15, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 15
;   LENGTH: 1006
;   TYPE: DNA
;   ORGANISM: SOYBEAN
US-09-247-373B-15

  Query Match             5.9%;  Score 29.4;  DB 4;  Length 1006;
```

```
  Best Local Similarity   60.8%;  Pred. No. 4.8;
  Matches   48;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      63 AGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAA 122
           ||||  ||      | | || |  | |||||  |||| |   ||||||||||| |||  |
Db     412 AGCTCAATTGCTTCCTCAACATTCTTCTCACGCTCTTTCTCATCAGCTGTGAAAACGGAT 353

Qy     123 GTAAAGGAAGCACGAAGAA 141
           |  | | |||||  || ||
Db     352 TTCCATGCAGCACCAAAAA 334


RESULT 10
US-09-296-715-15/c
; Sequence 15, Application US/09296715
; Patent No. 6171839
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/296,715
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1006 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SR1.PK0011.D6
US-09-296-715-15

  Query Match             5.9%;  Score 29.4;  DB 4;  Length 1006;
  Best Local Similarity   60.8%;  Pred. No. 4.8;
  Matches   48;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      63 AGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAA 122
           ||||  ||      | | || |  | |||||  |||| |   ||||||||||| |||  |
Db     412 AGCTCAATTGCTTCCTCAACATTCTTCTCACGCTCTTTCTCATCAGCTGTGAAAACGGAT 353

Qy     123 GTAAAGGAAGCACGAAGAA 141
           |  | | |||||  || ||
Db     352 TTCCATGCAGCACCAAAAA 334


RESULT 11
US-08-660-645A-11/c
; Sequence 11, Application US/08660645A
; Patent No. 6087152
;  GENERAL INFORMATION:
;    APPLICANT:  Hohmann, Hans-Peter
;    APPLICANT:  Pasamontes, Luis
;    APPLICANT:  Tessier, Michel
;    APPLICANT:  van Loon, Adolphus
;    TITLE OF INVENTION:  FERMENTATIVE CAROTENOID PRODUCTION
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/660,645A
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 95108888.9
;      FILING DATE:  09-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pokras, Bruce A.
;      REGISTRATION NUMBER:  32,748
;      REFERENCE/DOCKET NUMBER:  RAN 6002/170
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-5801
;      TELEFAX:  (201) 235-2363
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-660-645A-11

  Query Match             5.8%;  Score 29.2;  DB 3;  Length 726;
  Best Local Similarity   48.2%;  Pred. No. 4.6;
  Matches   82;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy     276 GGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGA 335
           ||| ||   ||||    ||| | ||||   ||    | ||||| || ||      | ||
Db     557 GGACGGTGCGGCAGCCAGGTACCGAAAACGAAGATCTGGATGGAAGCCAGAACAGCCGGA 498

Qy     336 AAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTT 395
           |     |      ||   ||   | ||  |   || |||||||||    ||  |  |
Db     497 ACCGGCCAGAAGATAACGTACATCCAACGGTCACCCAGGATCAGAGCGTAGGTGGTAACG 438

Qy     396 TCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAA 445
             |       ||| | |||    |   | || |||     |||||||||
Db     437 ATAACCGGCAGCAGCAGACCTTCACGCCAACCGAAGTAGGTGGAAACGAA 388


RESULT 12
US-09-298-718-11/c
; Sequence 11, Application US/09298718
; Patent No. 6124113
;  GENERAL INFORMATION:
;    APPLICANT:  Hohmann, Hans-Peter
;    APPLICANT:  Pasamontes, Luis
;    APPLICANT:  Tessier, Michel
;    APPLICANT:  van Loon, Adolphus
;    TITLE OF INVENTION:  FERMENTATIVE CAROTENOID PRODUCTION
```

```
;     NUMBER OF SEQUENCES:  47
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann-La Roche Inc.
;       STREET:  340 Kingsland Street
;       CITY:  Nutley
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07110
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/298,718
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/660,645
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pokras, Bruce A.
;       REGISTRATION NUMBER:  32,748
;       REFERENCE/DOCKET NUMBER:  RAN 6002/170
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (201) 235-5801
;       TELEFAX:  (201) 235-2363
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  726 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-298-718-11

  Query Match             5.8%;  Score 29.2;  DB 3;  Length 726;
  Best Local Similarity   48.2%;  Pred. No. 4.6;
  Matches   82;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy     276 GGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGA 335
           ||| ||   ||||    ||| | ||||   ||    | ||||| || ||      | ||
Db     557 GGACGGTGCGGCAGCCAGGTACCGAAAACGAAGATCTGGATGGAAGCCAGAACAGCCGGA 498

Qy     336 AAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTT 395
           |     |     ||   ||   | ||  |   || |||||||||    ||  |  |
Db     497 ACCGGCCAGAAGATAACGTACATCCAACGGTCACCCAGGATCAGAGCGTAGGTGGTAACG 438

Qy     396 TCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAA 445
             |       ||| | |||    |   | || |||     |||||||||
Db     437 ATAACCGGCAGCAGCAGACCTTCACGCCAACCGAAGTAGGTGGAAACGAA 388


RESULT 13
US-09-546-969-11/c
; Sequence 11, Application US/09546969
; Patent No. 6207409
;  GENERAL INFORMATION:
;     APPLICANT:  Hohmann, Hans-Peter
;     APPLICANT:  Pasamontes, Luis
;     APPLICANT:  Tessier, Michel
;     APPLICANT:  van Loon, Adolphus
;     TITLE OF INVENTION:  FERMENTATIVE CAROTENOID PRODUCTION
;     NUMBER OF SEQUENCES:  47
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann-La Roche Inc.
;       STREET:  340 Kingsland Street
;       CITY:  Nutley
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07110
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/546,969
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/660,645
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pokras, Bruce A.
;       REGISTRATION NUMBER:  32,748
;       REFERENCE/DOCKET NUMBER:  RAN 6002/170
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (201) 235-5801
;       TELEFAX:  (201) 235-2363
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  726 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-546-969-11

  Query Match             5.8%;  Score 29.2;  DB 4;  Length 726;
  Best Local Similarity   48.2%;  Pred. No. 4.6;
  Matches   82;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy     276 GGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGA 335
           ||| ||   ||||    ||| | ||||   ||    | ||||| || ||      | ||
Db     557 GGACGGTGCGGCAGCCAGGTACCGAAAACGAAGATCTGGATGGAAGCCAGAACAGCCGGA 498

Qy     336 AAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTT 395
           |     |     ||   ||   | ||  |   || |||||||||    ||  |  |
Db     497 ACCGGCCAGAAGATAACGTACATCCAACGGTCACCCAGGATCAGAGCGTAGGTGGTAACG 438

Qy     396 TCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAA 445
             |       ||| | |||    |   | || |||     |||||||||
Db     437 ATAACCGGCAGCAGCAGACCTTCACGCCAACCGAAGTAGGTGGAAACGAA 388


RESULT 14
US-08-980-832-28/c
; Sequence 28, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
;  APPLICANT: Pasamontes, Luis
;  APPLICANT: Tsygankov, Yuri
;  TITLE OF INVENTION: Improved Fermentative Carotenoid Production
;  FILE REFERENCE: Improved Fermentive Carotenoid
;  CURRENT APPLICATION NUMBER: US/08/980,832B
;  CURRENT FILING DATE: 1997-12-01
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 726
;   TYPE: DNA
;   ORGANISM: Alcaligenes PC-1
US-08-980-832-28

  Query Match             5.8%;  Score 29.2;  DB 4;  Length 726;
  Best Local Similarity   48.2%;  Pred. No. 4.6;
  Matches   82;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy     276 GGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGA 335
           ||| ||   ||||    ||| | ||||   ||    | ||||| || ||      | ||
Db     557 GGACGGTGCGGCAGCCAGGTACCGAAAACGAAGATCTGGATGGAAGCCAGAACAGCCGGA 498

Qy     336 AAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTT 395
           |     |     ||   ||   | ||  |   || |||||||||    ||  |  |
Db     497 ACCGGCCAGAAGATAACGTACATCCAACGGTCACCCAGGATCAGAGCGTAGGTGGTAACG 438
```

```
Qy     396 TCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAA 445
           |       ||| | |||     |   | || |||     |||||||||
Db     437 ATAACCGGCAGCAGCAGACCTTCACGCCAACCGAAGTAGGTGGAAACGAA 388
```

RESULT 15
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 5.8%; Score 29.2; DB 4; Length 4403765;
Best Local Similarity 59.8%; Pred. No. 68;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy     172 CTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGT 231
           ||| | | || ||| | |||  | | ||  |  ||| ||  |||| ||  ||||  ||
Db 4268434 CTCATAGATTGCCCCCGTGATCGGCTGAGCGTCGGTCACCGGCGACCAATACATGGCGCC 4268375

Qy     232 CTCGGCGAACCTTCCTTCGGGC 253
           ||  |||||  || | |   ||
Db 4268374 CTTAGCGAAGGTTGCATACCGC 4268353
```

Search completed: February 7, 2003, 09:05:32
Job time : 888.628 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.2903 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttttacaccatcaccaaa.....aggccacgtgatgatcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	204.4	40.9	6454	10	US-09-736-969A-1	Sequence 1, Appli
	2	85	17.0	4391	10	US-09-736-969A-7	Sequence 7, Appli
	3	71.4	14.3	7215	10	US-09-736-960-1	Sequence 1, Appli
	4	46	9.2	6372	9	US-09-736-968A-1	Sequence 1, Appli
	5	40.4	8.1	211	9	US-09-736-968A-77	Sequence 77, Appl
	6	34.6	6.9	6691	10	US-09-070-927A-88	Sequence 88, Appl
c	7	32.4	6.5	1314	9	US-09-738-626-795	Sequence 795, App
	8	32.4	6.5	3309400	9	US-09-738-626-1	Sequence 1, Appli
	9	31.8	6.4	659158	9	US-09-771-208-20	Sequence 20, Appl
c	10	31.6	6.3	14141	10	US-09-070-927A-394	Sequence 394, App
	11	31.4	6.3	4915	10	US-09-070-927A-125	Sequence 125, App
c	12	31.2	6.2	2000	9	US-09-938-842A-3143	Sequence 3143, Ap
c	13	30.8	6.2	417	10	US-09-960-352-12649	Sequence 12649, A
	14	30.6	6.1	32191	10	US-09-764-864-1678	Sequence 1678, Ap
c	15	30.4	6.1	1005	10	US-09-939-980-228	Sequence 228, App
c	16	30.4	6.1	326014	10	US-09-731-231A-3	Sequence 3, Appli
c	17	30.2	6.0	477	10	US-09-764-877-623	Sequence 623, App
	18	30.2	6.0	1972	10	US-09-765-272-203	Sequence 203, App
c	19	29.6	5.9	405	10	US-09-960-352-2213	Sequence 2213, Ap
c	20	29.4	5.9	251	9	US-09-796-692-9270	Sequence 9270, Ap
c	21	29.4	5.9	371	10	US-09-864-761-32600	Sequence 32600, A
c	22	29.4	5.9	528	9	US-09-796-692-3999	Sequence 3999, Ap
c	23	29.4	5.9	1686	9	US-10-114-893-37	Sequence 37, Appl
c	24	29.4	5.9	2707	10	US-09-917-800A-1420	Sequence 1420, Ap
	25	29.4	5.9	7925	10	US-09-070-927A-311	Sequence 311, App
c	26	29.4	5.9	9450	9	US-10-126-912-2	Sequence 2, Appli
c	27	29.4	5.9	9757	9	US-10-126-912-7	Sequence 7, Appli
c	28	29.4	5.9	9828	9	US-10-126-912-1	Sequence 1, Appli
c	29	29.2	5.8	726	9	US-09-920-923-28	Sequence 28, Appl
c	30	29.2	5.8	726	10	US-09-547-267-11	Sequence 11, Appl
	31	29.2	5.8	1685	10	US-09-827-998-19	Sequence 19, Appl
	32	29.2	5.8	1944	10	US-09-864-761-2825	Sequence 2825, Ap
	33	29.2	5.8	4158	10	US-09-827-998-15	Sequence 15, Appl
	34	29.2	5.8	5313	10	US-09-827-998-9	Sequence 9, Appli
	35	29.2	5.8	5376	10	US-09-827-998-2	Sequence 2, Appli
	36	29.2	5.8	5461	10	US-09-827-998-8	Sequence 8, Appli
	37	29.2	5.8	6719	10	US-09-827-998-1	Sequence 1, Appli
c	38	29	5.8	416	9	US-09-736-457-469	Sequence 469, App
c	39	29	5.8	416	9	US-09-902-941-469	Sequence 469, App
c	40	29	5.8	416	9	US-09-849-626-469	Sequence 469, App
	41	29	5.8	494	10	US-09-920-300A-446	Sequence 446, App
	42	29	5.8	494	12	US-10-033-528-446	Sequence 446, App
	43	29	5.8	590	10	US-09-864-761-9899	Sequence 9899, Ap
c	44	29	5.8	954	10	US-09-947-971-3	Sequence 3, Appli
	45	29	5.8	2586	10	US-09-804-551B-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-736-969A-1
; Sequence 1, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543

```
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 6454
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-4 cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (95)..(6121)
;   OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

  Query Match             40.9%;  Score 204.4;  DB 10;  Length 6454;
  Best Local Similarity   63.4%;  Pred. No. 8.1e-57;
  Matches  313;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           ||  |   ||  ||   ||||||| |||||  ||  ||||||||||||||||  |||  |  ||||  |
Db    2000 GTCTCGCATCACAACCAAAATCCAGAGTTCTATGATGAGATTAAAATTGAGCTTCCCATT 2059

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           ||  ||  |||  ||||  ||  ||   ||  |   ||||  ||  |   |||||  ||  |||||  |
Db    2060 CACCTACATCAAAAACATCATTTGCTTTTCACTTTTTATCATGTAAGTTGTGAAATTAAC 2119

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
           |   ||  ||||   ||  ||  |||    ||     |||||||  |  ||||||  |    |||||  |
Db    2120 ACAAAGGGAACAACCAAAAAGCAAGACACAGTTGAAACTCCAGTTGGGTTTGCCTGGGTA 2179

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||   |  ||||||||  ||  ||   |   |  |||    ||||||||   |  ||  ||  ||  ||  ||
Db    2180 CCTTTGCTGAAAGATGGTAGAATCATCACATTTGAGCAGCAGCTGCCAGTTTCCGCCAAT 2239

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           |||||  |  |||||  |        |  ||     |       |  |||||  |  |      |||
Db    2240 CTTCCCCCAGGCTACTTGAATCTGAATGATGCAGAATCAAGAAGGCAATGTAACGTGGAT 2299

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           |||||||||||||||  |  |||||  ||  ||||  ||    |   ||   |  |  |||||
Db    2300 ATTAAATGGGTAGATGGTGCAAAGCCTTTGTTGAAGTTTAAAAGCCACTTAGAATCTACC 2359

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCT 421
             |  ||  |||||  ||||  ||||  |  ||  ||  ||  |||||   |  ||  |||    |    |  ||
Db    2360 ATTTACACTCAAGATCTGCATGTGCACAAATTCTTCCATCATTGCCAGCTGATTCAGTCA 2419

Qy     422 GGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAA 481
           ||  |  |||    |    |    ||  ||   |  ||  ||   |  |||  ||  |||||||  |||||
Db    2420 GGCTCGAAAGAAGTTCCAGGGGAGCTCATTAAATATTTAAAGTGTTTGCATGCCATGGAG 2479

Qy     482 GGCCACGTGATGAT 495
             |||  ||  |||||
Db    2480 ATCCAAGTCATGAT 2493


RESULT 2
US-09-736-969A-7
; Sequence 7, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-4 Transmembrane Protein
;  FILE REFERENCE: 020054-000411US
;  CURRENT APPLICATION NUMBER: US/09/736,969A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 4391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human CLASP-4 cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (414)..(4058)
;   OTHER INFORMATION: human CLASP-4
US-09-736-969A-7

  Query Match             17.0%;  Score 85;  DB 10;  Length 4391;
  Best Local Similarity   52.4%;  Pred. No. 1.5e-17;
  Matches  187;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy     140 AAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGA 199
           |||  ||    ||  |  |  |  |||||  |   |||||  |  ||  |  |||||  |   ||
Db      78 AAGGCGGGCCCCGGTTGACCTCCGGTTGGGTTTGCCTGGGTCCCTTTGCTGAAGGTTGGT 137

Qy     200 AGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTT 259
           ||  |  |  |    | ||  ||   | |||||  ||  | |  |||||  ||||||   |
Db     138 AGATTCTTCCCTTTTGGGCGGCGGCTGCCGGTTTCCCCCATTCTTCCCCCGGGCTCCTTA 197

Qy     260 GGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGA 319
                 ||          |  ||||  |  |  |  |||  ||||||||||  | |||
Db     198 ATTCTGATTGATGGATAATCAAGGGGGCAATGTAATGTGGATTTTAAATGGGTGGTTGGT 257

Qy     320 GGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAG 379
           |  |||||  ||  ||   ||      |   |||  ||||  |  |  |   ||  |||  ||  |
Db     258 GTAAAGCCTTTGTGGAGGTTTAAAGGCCCCTTGGGTTCTCCCATTTCCCCTTAGGTTCTG 317

Qy     380 CATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGA 439
           |||||  |  ||  ||  |||||  |  ||  |  |   |     ||  |   |  |||   |   |
Db     318 CATTTGCCCAAATTCTTCCATCATTGCCGGCTGATTCGGTCAGACTCGAAAGAGGTTCCA 377

Qy     440 AACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAAGGCCACGTGATGATC 496
             |  ||  |  ||  |   |  |  |  |  |||||||  |||||   |||  ||  |||||
Db     378 GGGGGGCTCATTAAATTTTTAAGGTCTTTGCATGCCATGGAGATCCAAGTCTTGATC 434


RESULT 3
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 7215
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-5 cDNA
;   NAME/KEY: CDS
;   LOCATION: (112)..(6159)
;   OTHER INFORMATION: human CLASP-5
US-09-736-960-1

  Query Match             14.3%;  Score 71.4;  DB 10;  Length 7215;
  Best Local Similarity   51.1%;  Pred. No. 6e-13;
  Matches  208;  Conservative    0;  Mismatches  181;  Indels   18;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           |||  | |||||  |  |    || || |||||||| ||  | |||||  || | ||| ||
Db    1759 GTTACATACCATAATAAGTCTCCTGACTTTTATGAAGAAGTGAAAATTAAGCTCCCCGCT 1818

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
            ||||    | ||| ||||||||  || |||| |||| |||| |||||||| |   |
Db    1819 AAGCTCACAGTAAATCACCACCTCCTGTTCACCTTCTACCATATCAGCTGTCAGCA---- 1874

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
                       |||| |  | |    ||| ||||| |   | || || || |||||
Db    1875 --------------GAAGCAAGGAGCCTCCGTGGAAACTCTCCTGGGATATTCATGGCTG 1920

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||  | ||   | | | | |  |    ||  |     |     |||| ||  |   | |
Db    1921 CCAATTCTCTTAAATGAACGTCTTCAAACTGGATCCTACTGTCTCCCAGTTGCCTTGGAA 1980

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
                 |   | |      |   ||      || ||  | |       |   |||
Db    1981 AAATTGCCACCCAACTACTCCATGCATTCTGCTGAGAAAGTCCCATTACAGAATCCTCCC 2040

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           ||||| ||||  || |||   |||  | |  | || |||      ||    |||||| |
Db    2041 ATTAAGTGGGCTGAAGGACATAAGGGAGTATTTAATATTGAAGTGCAAGCTGTTTCTTCT 2100

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCA 408
           ||  | || |||||  | ||  |  | || || |||     ||| ||
Db    2101 GTACACACCCAGGACAACCACCTGGAGAAGTTCTTCACCCTCTGCCA 2147


RESULT 4
US-09-736-968A-1
; Sequence 1, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-7 Transmembrane Protein
;  FILE REFERENCE: 020054-000611US
;  CURRENT APPLICATION NUMBER: US/09/736,968A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 6372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-7 cDNA
;   NAME/KEY: CDS
;   LOCATION: (13)..(6156)
;   OTHER INFORMATION: human CLASP-7
US-09-736-968A-1

  Query Match             9.2%;  Score 46;  DB 9;  Length 6372;
  Best Local Similarity   63.6%;  Pred. No. 0.00013;
  Matches   70;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
```

```
          ||  |  ||||| || |   ||| || || || || ||| | ||  |  |  | ||  |
Db   1828 GTGGTCTACCATAACAAGTCCCCCGAGTTCTACGAGGAGTTCAAGCTGCATCTTCCAGCC 1887

Qy     62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTG 111
             ||   || || || |||||| || |||| |||| ||||||||||||
Db   1888 TGCGTGACAGAGAACCATCACCTGCTGTTCACCTTCTACCATGTCAGCTG 1937


RESULT 5
US-09-736-968A-77
; Sequence 77, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-7 Transmembrane Protein
;  FILE REFERENCE: 020054-000611US
;  CURRENT APPLICATION NUMBER: US/09/736,968A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
;   LENGTH: 211
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: exon 88470-88680
US-09-736-968A-77

  Query Match             8.1%;  Score 40.4;  DB 9;  Length 211;
  Best Local Similarity   65.6%;  Pred. No. 0.0013;
  Matches   59;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      22 CCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCA 81
           ||| || || || || ||| | ||  |  |  | ||  |     ||   || || || ||
Db      38 CCCCGAGTTCTACGAGGAGTTCAAGCTGCATCTTCCAGCCTGCGTGACAGAGAACCATCA 97

Qy      82 CCTGTTGCTCACATTCTTCCATGTCAGCTG 111
           |||| || |||| |||| ||||||||||||
Db      98 CCTGCTGTTCACCTTCTACCATGTCAGCTG 127


RESULT 6
US-09-070-927A-88
; Sequence 88, Application US/09070927A
; Patent No. US20020120116A1
;  GENERAL INFORMATION:
;       APPLICANT: Charles A. Kunsch
;                  Patrick J. Dillon
;                  Steven Barash
;       TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;       NUMBER OF SEQUENCES: 982
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/070,927A
;            FILING DATE: 04-May-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/046,655
;            FILING DATE: 1997-05-16
;            APPLICATION NUMBER: 60/044,031
;            FILING DATE: 1997-05-06
;            APPLICATION NUMBER: 60/066,009
;            FILING DATE: 1997-11-14
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kenley K. Hoover
;            REGISTRATION NUMBER: 40,302
;            REFERENCE/DOCKET NUMBER: PB369
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (301) 309-8504
;            TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 88:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 6691 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 88:
US-09-070-927A-88

  Query Match             6.9%;  Score 34.6;  DB 10;  Length 6691;
  Best Local Similarity   54.3%;  Pred. No. 0.78;
  Matches   70;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy      12 ATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACTCAGCTGCATG 71
           | ||||||||   | | |   |  |  |||| || |||  ||||  ||   ||||||
Db    1656 AACACCAAAAGAAATAGTCAAAGAACTAGATCAATATATTGTTGGACAGCAAGCTGCCAA 1715

Qy      72 AAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAAGTAAAGGAA 131
           |||| ||    | || |||     |   |  | ||  |  |  ||| |  |  |||  |
Db    1716 AAAATCAGTGGCGGTAGCTTTACGTAACCGCTATCGTCGCTTGCAATTAGAAGAAAATAT 1775

Qy     132 GCACGAAGA 140
           |||  ||||
Db    1776 GCAACAAGA 1784


RESULT 7
US-09-738-626-795/c
; Sequence 795, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
```

```
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 795
;   LENGTH: 1314
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-795

  Query Match             6.5%;  Score 32.4;  DB 9;  Length 1314;
  Best Local Similarity   54.1%;  Pred. No. 1.6;
  Matches   66;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy      40 GATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTT 99
           ||||  |   ||  ||   |   ||||  |     |||  ||  |||||   |||   |||
Db     892 GATTTAGTGAGTCTTTGACCGGCAGCAACGCATTAAGAGCCTGCTGTTTTGCACGCGCTT 833

Qy     100 CCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAAC 159
           |     | |  ||| ||||  |||| |   || || |||| |  || |||   | |  ||
Db     832 CTTCCACGGTCGTGCCAACAAAAGTCACCAAACCAGGAAGCACCGGCATGGGCTCGCGAC 773

Qy     160 CC 161
           ||
Db     772 CC 771


RESULT 8
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

  Query Match             6.5%;  Score 32.4;  DB 9;  Length 3309400;
  Best Local Similarity   54.1%;  Pred. No. 91;
  Matches   66;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy      40 GATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTT 99
           ||||  |   ||  ||   |   ||||  |     |||  ||  |||||   |||   |||
Db 737638 GATTTAGTGAGTCTTTGACCGGCAGCAACGCATTAAGAGCCTGCTGTTTTGCACGCGCTT 737697

Qy     100 CCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAAC 159
           |     | |  ||| ||||  |||| |   || || |||| |  || |||   | |  ||
Db 737698 CTTCCACGGTCGTGCCAACAAAAGTCACCAAACCAGGAAGCACCGGCATGGGCTCGCGAC 737757

Qy     160 CC 161
           ||
Db 737758 CC 737759


RESULT 9
US-09-771-208-20
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
;  APPLICANT: MEDRANO, JUAN
;  APPLICANT:  BRADFORD, ERIC
;  APPLICANT:  HORVAT, SIMON
;  TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
;  FILE REFERENCE: 407T-923710US
;  CURRENT APPLICATION NUMBER: US/09/771,208
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER: US 08/999,477
;  PRIOR FILING DATE: 1997-12-29
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 659158
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (123459)..(123478)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (602466)..(602485)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (546998)..(547017)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (494715)..(494814)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (390986)..(391005)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (346860)..(346823)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (317174)..(317193)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (280353)..(280373)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (271829)..(271848)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (183872)..(183891)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (170625)..(170645)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (132680)..(132700)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20

  Query Match             6.4%;  Score 31.8;  DB 9;  Length 659158;
  Best Local Similarity   67.2%;  Pred. No. 86;
  Matches   45;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      53 TTGCCCACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGT 112
           ||   || ||  ||   ||| ||| |  ||||||| ||| | | |||  || |  |||||
Db  138799 TTTGACATTCTTCTCACTGAGAAGTAGGACCTGTTTCTCTCCTCCTTGGATCTGGGCTGT 138858

Qy     113 GACAACT 119
           |||||||
Db  138859 GACAACT 138865


RESULT 10
US-09-070-927A-394/c
; Sequence 394, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 394:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14141 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 394:
US-09-070-927A-394

  Query Match             6.3%;  Score 31.6;  DB 10;  Length 14141;
  Best Local Similarity   56.9%;  Pred. No. 12;
  Matches   58;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy      66 TGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAAGTA 125
           ||| | ||  ||  ||| ||||| |||  ||||| || | | ||| |  |  |  || ||
Db    7335 TGCTTAAACGGCCACACTTGTTGTTCATCTTCTTTCAAGACCGCTTTTCCTTCAAAAATA 7276

Qy     126 AAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTG 167
           |||   |  || ||| |      |    |||| | || | |
Db    7275 TAGGTTCCTGGACGAACATCTTCTTTCCCGAAATCTAAATGG 7234


RESULT 11
US-09-070-927A-125
; Sequence 125, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 125:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 4915 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 125:
US-09-070-927A-125

  Query Match             6.3%;  Score 31.4;  DB 10;  Length 4915;
  Best Local Similarity   57.7%;  Pred. No. 7.4;
  Matches   56;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy     299 GAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCT 358
           ||||  || |   |   |    | ||   ||  ||  | | |   | |||| || ||||
Db     235 GAAACAAACTCACTTTTTTACTGAAATTTACATCTTGATAATAATAATCATTTGTTTTCA 294

Qy     359 ACAGTGTATACTCAGGATCAGCATTTACATAATTTTT 395
             | |  |||||| |||| | ||| || ||||| |||
Db     295 TAACTACATACTCTGGATAATCATATAGATAATCTTT 331
```

```
RESULT 12
US-09-938-842A-3143/c
; Sequence 3143, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3143
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-3143

  Query Match             6.2%;  Score 31.2;  DB 9;  Length 2000;
  Best Local Similarity   50.7%;  Pred. No. 5.1;
  Matches   75;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy      17 CAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACTCAGCTGCATGAAAAG 76
           ||||| |||  |||   || |   || | |  | | ||   |||| |  |  | |||| |
Db     893 CAAAAACCACGATTCCTTGTTACTATAATATAATACTTCTTCACTAAAATCAAAGAAATG 834

Qy      77 CACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACG 136
           ||    |    || ||  || ||| |   | | | || | ||| |||   |    |
Db     833 GGAAACTACCTTGGGACGTTTATCAATGGCGAATTTAAGAAGTAAAGCAAAACCAATTTG 774

Qy     137 AAGAAGAGGGATGTCGTTGAAACCCAAG 164
           ||||         || || ||| |  ||
Db     773 AAGAGATTCCCAATCATTCAAATCACAG 746


RESULT 13
US-09-960-352-12649/c
; Sequence 12649, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12649
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 54-LIB3058-010-Q1-K1-F6
US-09-960-352-12649

  Query Match             6.2%;  Score 30.8;  DB 10;  Length 417;
  Best Local Similarity   49.4%;  Pred. No. 2.8;
  Matches   80;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy     323 AAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCAT 382
           |||| |  | | |||||   || |      |  |  |    | || |  ||     || |
Db     162 AAGCTATAGTTAAAAATGAACAGTTTATATGAATGAAAGCAGCATGCATAGTCAAGGCTT 103

Qy     383 TTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAAC 442
           || |||||  |||||     ||| || || |   |      | |||||| | |   |
Db     102 TTTCATAACATTTTCATTAGCTGACAAAACATTAAGGTAAAAACCCAAGCTTAATATATA 43

Qy     443 GAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGAAGGC 484
           || |  | ||  |      | | | ||  |||||| | | |
Db      42 TAATTCATTAAAAAAAAAGGGCGGCCGCTCGCGATGTAGGAC 1


RESULT 14
US-09-764-864-1678
; Sequence 1678, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1678
;   LENGTH: 32191
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-864-1678

  Query Match             6.1%;  Score 30.6;  DB 10;  Length 32191;
  Best Local Similarity   58.1%;  Pred. No. 40;
  Matches   54;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy     382 TTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAA 441
           ||||| || |||| |   | | | ||| ||   | ||  || |||  ||   ||   |||
Db    3777 TTTACTTATTTTTGTAATGGATTTTCATAATGGCTAACATGTAGCTGAATAATTCTAAAA 3836

Qy     442 CGAACTTGTAAAGTACCTTAAGAGTCTGCATGC 474
           | |  || |||||  ||| | | || || |
Db    3837 TACATTATTATAGTACAATAATATTTTGAATTC 3869


RESULT 15
US-09-939-980-228/c
; Sequence 228, Application US/09939980
; Patent No. US20020082234A1
;   GENERAL INFORMATION:
;        APPLICANT: Black, Michael
;                   Burnham, Martin
;                   Hodgson, John
;                   Knowles, David
;                   Lonetto, Michael
;                   Nicholas, Richard
;                   Pratt, Julie
;                   Reichard, Richard
;                   Rosenberg, Martin
;                   Ward, Judith
;        TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
;                            Polypeptides and Their Uses
;        NUMBER OF SEQUENCES: 534
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SmithKline Beecham Corporation
;             STREET: 709 Swedeland Road
;             CITY: King of Prussia
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19406-0939
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
```

```
;              APPLICATION NUMBER: US/09/939,980
;              FILING DATE: 27-Aug-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/936,165
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Gimmi, Edward R
;              REGISTRATION NUMBER: 38,891
;              REFERENCE/DOCKET NUMBER: P50549
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 610-270-4478
;              TELEFAX: 610-270-5090
;              TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 228:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1005 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Genomic DNA
;         SEQUENCE DESCRIPTION: SEQ ID NO: 228:
US-09-939-980-228

  Query Match             6.1%;  Score 30.4;  DB 10;  Length 1005;
  Best Local Similarity   62.5%;  Pred. No. 6.3;
  Matches   65;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGA---TGAGATTAAAATAGAGTTGCCC 58
           ||| | ||||| ||  || ||| | ||||| ||||   | ||  |  |||  || | | |
Db     439 GTTGTCCACCAGCAATAATACCTATAATTTCATGAGGGTCAGTATTGAATGTAGGTACAC 380

Qy      59 ACTCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCA 102
           |||| |||||||  ||  ||||   | || | |   | | | ||
Db     379 ACTCCGCTGCATCTAAGACACCCAATCTTCCACTTGTACCTGCA 336

Search completed: February  7, 2003, 09:14:04
Job time : 1286.29 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 ; Search time 1008.32 Seconds
(without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_1_500
Perfect score: 500
Sequence: 1 agttttacaccatcaccaaa.....aggccacgtgatgatcgcct 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	457.6	91.5	800	12	BE783911	BE783911 601470964
2	359.8	72.0	465	12	BE938152	BE938152 MR3-TN004
3	347.2	69.4	465	12	BE938144	BE938144 MR3-TN004
4	308	61.6	891	14	BQ919716	BQ919716 AGENCOURT
5	166	33.2	412	10	BB672508	BB672508 BB672508
6	166	33.2	441	10	BB819621	BB819621 BB819621
7	148.2	29.6	245	12	BF948041	BF948041 MR3-NN021
8	144.4	28.9	538	12	BF472403	BF472403 UI-M-BH3-
9	134.8	27.0	543	10	BB763030	BB763030 BB763030
10	117	23.4	345	10	BB222768	BB222768 BB222768
11	106	21.2	330	10	BB207661	BB207661 BB207661
12	100.2	20.0	285	10	BB225426	BB225426 BB225426
13	94.4	18.9	274	10	BB177025	BB177025 BB177025
14	90.4	18.1	314	10	BB219663	BB219663 BB219663
15	89.8	18.0	249	10	BB538596	BB538596 BB538596
16	83.6	16.7	212	10	AV344231	AV344231 AV344231
17	80.6	16.1	255	10	BB200610	BB200610 BB200610
18	78.2	15.6	259	10	BB241632	BB241632 BB241632
19	77.4	15.5	219	10	BB184661	BB184661 BB184661
c 20	76.6	15.3	917	17	CNS03GUC	AL243453 Tetraodon
21	74	14.8	447	12	BG019353	BG019353 daa78g06.
22	73.2	14.6	239	10	BB256604	BB256604 BB256604
23	57	11.4	669	13	BI961883	BI961883 MONO1_7_H
24	55.6	11.1	749	13	BI828723	BI828723 603074748
25	53.4	10.7	147	14	BQ339280	BQ339280 MR3-NN021
26	51.6	10.3	271	13	BM029853	BM029853 488237 MA
c 27	49.2	9.8	534	12	BE945462	BE945462 UI-M-BH3-
28	49.2	9.8	656	13	BJ395781	BJ395781 BJ395781
29	46	9.2	909	14	BQ730710	BQ730710 AGENCOURT
30	45.6	9.1	695	14	C88441	C88441 C88441 Carp
31	42.4	8.5	506	9	AA391952	AA391952 LD10883.5
c 32	40.8	8.2	254	17	AZ632934	AZ632934 1M0487B21
33	39	7.8	657	13	BG914864	BG914864 602813690
34	39	7.8	1507	11	AK010755	AK010755 Mus muscu
35	38	7.6	522	10	BE667330	BE667330 151645 MA
36	37.6	7.5	631	13	BI507332	BI507332 BB170003A
37	37	7.4	635	9	AA892774	AA892774 EST196577
38	37	7.4	642	14	BQ206156	BQ206156 UI-R-EP0-
c 39	37	7.4	753	14	BQ201592	BQ201592 UI-R-DZ1-
c 40	37	7.4	836	17	CNS009OK	AL053937 Drosophil
c 41	36.6	7.3	802	17	CNS0178I	AL107724 Drosophil
c 42	36.2	7.2	350	10	AW784142	AW784142 NXNV_103_
43	36	7.2	595	9	AA695066	AA695066 GM02069.5
44	36	7.2	864	13	BI558960	BI558960 603241321
45	35.4	7.1	995	17	CNS02IU2	AL199379 Tetraodon

ALIGNMENTS

RESULT 1
BE783911
LOCUS BE783911 800 bp mRNA linear EST 20-OCT-2000
DEFINITION 601470964F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3874046 5', mRNA sequence.
ACCESSION BE783911
VERSION BE783911.1 GI:10205109
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9630 row: p column: 15
High quality sequence stop: 629.
FEATURES Location/Qualifiers
source 1. .800

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3874046"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      227 a    198 c    185 g    190 t
ORIGIN

  Query Match             91.5%;  Score 457.6;  DB 12;  Length 800;
  Best Local Similarity   98.6%;  Pred. No. 6.9e-133;
  Matches  493;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy       4 TTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAA-TAGAGTTGCCCACTC 62
           |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db      20 TTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAGCTAGAGTTGCCCACTC 79

Qy      63 AGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAAC-TCA 121
           ||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db      80 AGCTGCATGAAAAGCACCACCTGTAGCTCACATTCTTCCATGTCAGCTGTGACAACGTCA 139

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 199

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 259

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     260 CTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 319

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAA-AATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     320 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAACAATTTCCACTCATCTGGTTTCTAC 379

Qy     361 AGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAAACCGAATC 420
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     380 AGTGTATACTCAGGATCAGCATTTACATAATTTTCTCCAGTACTGTCAGAAAACCGAATC 439

Qy     421 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 TGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCATGCGATGGA 499

Qy     481 AGGCCACGTGATGATCGCCT 500
           ||||||||||||||||||||
Db     500 AGGCCACGTGATGATCGCCT 519


RESULT 2
BE938152
LOCUS       BE938152                 465 bp    mRNA    linear   EST 02-OCT-2000
DEFINITION  MR3-TN0048-280800-001-g08 TN0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE938152
VERSION     BE938152.1  GI:10465154
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-TN0048-280
            800-001-g08&t3=2000-08-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 465.
FEATURES             Location/Qualifiers
     source          1. .465
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="TN0048"
                     /dev_stage="Adult"
                     /note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      127 a    114 c    115 g    109 t
ORIGIN

  Query Match             72.0%;  Score 359.8;  DB 12;  Length 465;
  Best Local Similarity   99.4%;  Pred. No. 3.5e-102;
  Matches  361;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     103 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 162

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     163 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 222

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     223 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 282

Qy     181 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     283 TCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAA 342

Qy     241 CCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 300
           ||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     343 CCTTCCTCCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGA 402

Qy     301 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     403 AATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTAC 462

Qy     361 AGT 363
           |||
Db     463 AGT 465


RESULT 3
BE938144
LOCUS       BE938144                 465 bp    mRNA    linear   EST 02-OCT-2000
DEFINITION  MR3-TN0048-280800-001-e05 TN0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE938144
VERSION     BE938144.1  GI:10465135
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-TN0048-280
            800-001-e05&t3=2000-08-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 465.
FEATURES             Location/Qualifiers
     source          1. .465
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="TN0048"
                     /dev_stage="Adult"
                     /note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      129 a    112 c    116 g    108 t
ORIGIN

  Query Match             69.4%;  Score 347.2;  DB 12;  Length 465;
  Best Local Similarity   98.9%;  Pred. No. 3.3e-98;
  Matches  360;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       1 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 AGTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCAC 161

Qy      61 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TCAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTC 221

Qy     121 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCT 180
          |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||
Db     222 AAGTAAAGGAAGCACGAAGAAGAGGGATGTCGATGAAACCCAAGTGGGCTACTCCTGGCT 281

Qy     181 TCCCCTCCTGAAAGACGGAAGGG-TGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGA 239
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     282 TCCCCTCCTGAAAGACGGAAGGGTTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGA 341

Qy     240 ACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGG 299
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     342 ACCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGG 401

Qy     300 AAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTA 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 AAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTA 461

Qy     360 CAGT 363
           ||||
Db     462 CAGT 465


RESULT 4
BQ919716
LOCUS       BQ919716                 891 bp    mRNA    linear   EST 20-AUG-2002
DEFINITION  AGENCOURT_8858966 NCI_CGAP_Mam2 Mus musculus cDNA clone
            IMAGE:6441934 5', mRNA sequence.
ACCESSION   BQ919716
VERSION     BQ919716.1  GI:22334414
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 891)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13965  row: o  column: 23
            High quality sequence stop: 412.
FEATURES             Location/Qualifiers
     source          1. .891
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6441934"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      243 a    209 c    236 g    201 t      2 others
ORIGIN

  Query Match             61.6%;  Score 308;  DB 14;  Length 891;
  Best Local Similarity   87.1%;  Pred. No. 1e-85;
  Matches  338;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy     113 GACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTAC 172
           || ||||| |  |||||||||||||||||||| ||| |  || ||||| || |||||||
Db       1 GATAACTCCACCAAAGGAAGCACGAAGAAGAAGGACGCTGTGGAAACGCAGGTTGGCTTT 60

Qy     173 TCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTC 232
           |||||||| || |||||||||||  |||||||| |||| || |||||||||||||| |||
Db      61 TCCTGGCTGCCTCTCCTGAAAGATGGAAGGGTGTTGACGAGTGAGCAGCACATCCCCGTC 120

Qy     233 TCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTAT 292
           ||||| ||||| || || ||||| || |||||||| ||||| || ||||||||||||||||
Db     121 TCGGCTAACCTGCCATCTGGCTACCTCGGCTACCAGGAGCTCGGCATGGGCAGGCATTAT 180

Qy     293 GGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTG 352
           ||||| ||  |||| ||||| || |||||||||||||| |||| || ||||||||||||
Db     181 GGTCCAGAGGTTAAGTGGGTGGAAGGAGGCAAGCCACTGTTGAAGATCTCCACTCATCTG 240

Qy     353 GTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTCAGAAA 412
           ||||| |||||||| |||||||||| |||||||||||||||||||| ||||||||||||
Db     241 GTTTCCACAGTGTACACTCAGGATCAACATTTACATAATTTTTTCCAATACTGTCAGAAA 300
```

```
Qy     413 ACCGAATCTGGAGCCCAAGCCTTAGGAAACGAACTTGTAAAGTACCTTAAGAGTCTGCAT 472
           || |||||||||||||||||||| ||| |  ||||| ||||| ||||||||||||||||||
Db     301 ACGGAATCTGGAGCCCAAGCCTCAGGGAGTGAACTAGTAAAATACCTTAAGAGTCTGCAT 360

Qy     473 GCGATGGAAGGCCACGTGATGATCGCCT 500
           |||||||||||||| ||||||||||||
Db     361 GCGATGGAAGGCCATGTGATGATCGCCT 388


RESULT 5
BB672508
LOCUS       BB672508                412 bp    mRNA    linear   EST 05-OCT-2001
DEFINITION  BB672508 RIKEN full-length enriched, adult male brain Mus musculus
            cDNA clone 3526405E11 3', mRNA sequence.
ACCESSION   BB672508
VERSION     BB672508.1  GI:15971729
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="3526405E11"
                     /clone_lib="RIKEN full-length enriched, adult male brain"
                     /sex="male"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I"
BASE COUNT      130 a    100 c     76 g    106 t
ORIGIN

  Query Match             33.2%;  Score 166;  DB 10;  Length 412;
  Best Local Similarity   63.1%;  Pred. No. 3.8e-41;
  Matches  256;  Conservative    0;  Mismatches  150;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
             ||  ||||||||   ||||| ||||| || | ||||||| | ||||| ||  | || ||
Db       7 GTGCTACACCATTCTCAAAATCCAGATTTCTCTGATGAGGTGAAAATTGAACTACCAACA 66

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           || || |||| ||| || ||||| ||| || | || | |||  |||  ||||||| |
Db      67 CAACTCCATGGAAAACATCACCTTTTGTTCTCCTTTTACCACATCACATGTGACATCAAT 126

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
             ||||  |   | || |||| || |   | || ||   ||| |||||  | ||||||
Db     127 GCCAAAGCCAATGCCAAAAAGAAAGAGGCTTTGGAGACATCAGTGGGCTATGCATGGCTT 186

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           || ||  ||||| | |      | |   |    ||| |  ||||||| |  |  ||| |
Db     187 CCTCTGATGAAACATGATCAAATAGCTTCTCAGGAGTACAACATCCCAATAGCAACGACC 246

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           || ||| |    ||| |  ||   ||||| | |   |  | |||  |||    ||
Db     247 CTGCCTCCTAATTATTTAAGCATTCAAGATCCTACAAGTGCAAAGCATGGTGGAAGTGAC 306

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           |||||||||| ||||| ||||| || ||  | ||| | |||||   | | || || |||
Db     307 ATTAAATGGGTCGATGGTGGCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACA 366

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTC 407
           ||| | |||||||| |  ||| || ||   |||||||   | || |
Db     367 GTGAACACTCAGGACCCACATGTAAATGCATTTTTCCGTCAGTGCC 412


RESULT 6
BB819621
LOCUS       BB819621                441 bp    mRNA    linear   EST 19-NOV-2001
DEFINITION  BB819621 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus
            musculus cDNA clone G730050N08 3', mRNA sequence.
ACCESSION   BB819621
VERSION     BB819621.1  GI:16992250
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
```

```
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES              Location/Qualifiers
     source           1. .441
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="G730050N08"
                      /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
                      cDNA"
                      /tissue_type="lung"
                      /cell_line="RCB-0558 LLC"
                      /note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164),
                      (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
                      (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
                      (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16
                      melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
                      ), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
                      (cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
                      (tissue_type=bladder, cell_line=RCB-0544 MBT-2),
                      (tissue_type=bone marrow, cell_type=stroma cell,
                      cell_line= CRL-2028 SR-4987), (tissue_type=colon,
                      cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
                      cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
                      cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
                      cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
                      ), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
                      BC3H1)"
BASE COUNT      133 a    113 c     81 g    114 t
ORIGIN

  Query Match             33.2%;  Score 166;  DB 10;  Length 441;
  Best Local Similarity   63.1%;  Pred. No. 4e-41;
  Matches  256;  Conservative    0;  Mismatches  150;  Indels    0;  Gaps    0;

Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           ||  ||||||||   ||||| ||||| || | ||||||| | ||||| ||  | || ||
Db      36 GTGCTACACCATTCTCAAAATCCAGATTTCTCTGATGAGGTGAAAATTGAACTACCAACA 95

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           || || |||| ||| || ||||| ||| || | || | |||  |||  ||||||| |
Db      96 CAACTCCATGGAAAACATCACCTTTTGTTCTCCTTTTACCACATCACATGTGACATCAAT 155

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
```

```
           ||||  |   | || ||||  || |   | || ||   ||| |||||  | |||||||
Db     156 GCCAAAGCCAATGCCAAAAAGAAAGAGGCTTTGGAGACATCAGTGGGCTATGCATGGCTT 215

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           || ||  ||||| | |       | |   |    ||| |  |||||||  |  |  ||| |
Db     216 CCTCTGATGAAACATGATCAAATAGCTTCTCAGGAGTACAACATCCCAATAGCAACGACC 275

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAA 301
           || ||| |    ||| |  ||   ||||| | |   |  |  | ||||  |||    ||
Db     276 CTGCCTCCTAATTATTTAAGCATTCAAGATCCTACAAGTGCAAAGCATGGTGGAAGTGAC 335

Qy     302 ATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACA 361
           |||||||||||| ||||| ||||| || ||  | ||| | |||||   | | || || |||
Db     336 ATTAAATGGGTCGATGGTGGCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACA 395

Qy     362 GTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTC 407
           ||| | |||||||| |  || || ||   |||||||   | || |
Db     396 GTGAACACTCAGGACCCACATGTAAATGCATTTTTCCGTCAGTGCC 441
```

```
RESULT 7
BF948041
LOCUS       BF948041                 245 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  MR3-NN0215-311000-007-b01 NN0215 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF948041
VERSION     BF948041.1  GI:12365316
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 245)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-NN0215-
            311000-007-b01&t3=2000-10-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 77.
FEATURES              Location/Qualifiers
     source           1. .245
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="NN0215"
                      /dev_stage="Adult"
                      /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
BASE COUNT       62 a     52 c     76 g     55 t
ORIGIN

  Query Match             29.6%;  Score 148.2;  DB 12;  Length 245;
```

```
  Best Local Similarity   89.9%;  Pred. No. 1.2e-35;
  Matches  204;  Conservative    0;  Mismatches   18;  Indels    5;  Gaps    4;

Qy     195 ACGGAAGGGTGGTGACAA-GCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGC 253
           |||||||||||||||| || |||||||| |||||||||||||||||||| ||||||||||||
Db      19 ACGGAAGGGTGGTGAGAAGGCGAGCAGTACATCCCGGTCTCGGCGAAGCTTCCTTCGGGC 78

Qy     254 TATCT-TGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGT 312
           ||||  ||| ||||| ||||||||||||||||||||||||||||||||||||| ||||||||
Db      79 GATCTGTGGATACCAGGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATGAAATGGGT 138

Qy     313 AGATGGAGGCAAGCCACTGCTGAAAATTTC-CACTCATCTGGTTTCTACAGTGTATACTC 371
           |||||||||| |||||||||||||| || ||||||||||||||||||||||||||||||
Db     139 AGATGGAGGCAGGCCACTGCTGAAAATGTCGCACTCATCTGGTTTCTACAGTGTATACTC 198

Qy     372 AGGATCAGCATTTACATAATTTT--TTCCAGTACTGTCAGAAAACCG 416
           |||||||| ||||||   | | |  | |  ||||||||||||||||
Db     199 AGGATCAGGATTTACGATAATATGGTCCATGTACTGTCAGAAAACCG 245


RESULT 8
BF472403
LOCUS       BF472403                538 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  UI-M-BH3-awc-c-08-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
            UI-M-BH3-awc-c-08-0-UI 5', mRNA sequence.
ACCESSION   BF472403
VERSION     BF472403.1  GI:11541586
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
FEATURES            Location/Qualifiers
     source         1. .538
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="UI-M-BH3-awc-c-08-0-UI"
                    /clone_lib="NIH_BMAP_M_S4"
                    /dev_stage="27-32 days"
                    /lab_host="DH10B (Life Technologies)"
                    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                    polylinker; Site_1: Not I; Site_2: Eco RI; The
                    NIH_BMAP_M_S4 library is a subtracted library of a series,
                    ultimately  derived from a mixture of individually tagged
                    normalized libraries from  ten regions of the mouse brain
                    (cerebellum, brain stems, olfactory  bulbs, hypothalamus,
                    cortex, amygdala, basal ganglia, pineal gland,  striatum,
                    hipoccampus) after a series of subtractions to reduce the
                    representation of cDNAs from which ESTs had already been
                    generated. The following serially subtracted libraries
                    were generated in this process: NIH_BMAP_M_S4,
                    NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
                    NIH_BMAP_M_S2, NIH_BMAP_M_S1.  The subtracted library
                    (NIH_BMAP_M_S4) was constructed as follows:  PCRamplified
                    cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
                    NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
                    was used as  a driver in a hybridization with a pool of
                    the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
                    libraries in the form of  single-stranded circles.  The
                    remaining single-stranded circles  (subtracted library)
                    was purified by hydroxyapatite column chromatography,
                    converted to double-stranded circles and electroporated
                    into DH10B  bacteria (LifeTechnologies) to generate the
                    NIH_BMAP_M_S4 library.  This procedure has been previously
                    described (Bonaldo, Lennon and Soares, Genome Research
                    6:791-806, 1996)"
BASE COUNT      150 a    122 c    123 g    143 t
ORIGIN

  Query Match           28.9%;  Score 144.4;  DB 12;  Length 538;
  Best Local Similarity   90.6%;  Pred. No. 2.8e-34;
  Matches  154;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     331 GCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAA 390
           | |||| || |||||||||||||||| |||||||| |||||||||| ||||||||||
Db      13 GTTGAAGATCTCCACTCATCTGGTTTCCACAGTGTACACTCAGGATCAACATTTACATAA 72

Qy     391 TTTTTTCCAGTACTGTCAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAACTTGT 450
           |||||||| |||||||||||||| |||||||||||||||||| ||  |  ||||| ||
Db      73 TTTTTTCCAATACTGTCAGAAAACGGAATCTGGAGCCCAAGCCTCAGCGAGTGAACTAGT 132

Qy     451 AAAGTACCTTAAGAGTCTGCATGCGATGGAAGGCCACGTGATGATCGCCT 500
           ||| |||||||||||||||||||||||||||||| ||||||||||||
Db     133 AAAATACCTTAAGAGTCTGCATGCGATGGAAGGCCATGTGATGATCGCCT 182


RESULT 9
BB763030
LOCUS       BB763030                543 bp    mRNA    linear   EST 17-OCT-2001
DEFINITION  BB763030 RIKEN full-length enriched, B16 F10Y cells Mus musculus
            cDNA clone G370018M23 3', mRNA sequence.
ACCESSION   BB763030
VERSION     BB763030.1  GI:16207943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
```

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara ,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

```
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G370018M23"
                     /clone_lib="RIKEN full-length enriched, B16 F10Y cells"
                     /cell_type="B16 F10Y cells"
                     /note="pooled tissues ; (tissue_type=cerebellum,
                     dev_stage=16 days neonate, sex=mixed),
                     (tissue_type=cerebellum, dev_stage=0 day neonate,
                     sex=mixed), (tissue_type=hippocampus, dev_stage=adult,
                     sex=male), (tissue_type=whole body, dev_stage=9 days
                     embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days
                     embryo, sex=mixed)"
BASE COUNT      164 a    137 c    101 g    141 t
ORIGIN
```

Query Match 27.0%; Score 134.8; DB 10; Length 543;
Best Local Similarity 63.0%; Pred. No. 3e-31;
Matches 257; Conservative 0; Mismatches 147; Indels 4; Gaps 3;

```
Qy       2 GTTTTACACCATCACCAAAACCCAGAATTTTATGATGAGATTAAAATAGAGTTGCCCACT 61
           ||  ||||||||   ||||| ||||| || | ||||||| | ||||| ||  | || ||
Db     138 GTGGTACACCATTCTCAAAATCCAGATTTCTGTGATGAGGTGAAAATTGAACTACCAACA 197

Qy      62 CAGCTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCA 121
           || || |||||||| || ||||| ||| || | || | |||  |||  ||||||| |
Db     198 CAACTCCATGAAAA-CATCACCTTTTGTTCTCCTTTTACCACATCACATGTGACATCAAT 256

Qy     122 AGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTT 181
               ||||  |   | || ||||  || |   | || ||   ||| |||||  | ||||||
Db     257 GCCAAAGCCAATGCCAAAAAGAAAGAGGCTTTGGAGACATCAGTGGGCTATGCATGGCTT 316

Qy     182 CCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAAC 241
           || ||  ||||| | |      | |   |    ||| |  |||||||  |  |  | | |
Db     317 CCTCTGATGAAACATGATCAAATAGCTTCTCAGGAGTACAACATCCCAATACCAACCACC 376

Qy     242 CTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGC--AGGCATTATGGTCCGG 299
           || ||| |    ||| |  ||   ||||| | | | | |  |  ||||||  |||    |
Db     377 CTCCCTCCTAATTATTTAAGCATTCAAGATCCTAGCAAGTCCAAAGGCATGGTGGAAGTG 436

Qy     300 AAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTA 359
           | ||||||||| | ||| | ||||| || ||  | ||| | |||||   | | || || |
Db     437 ACATTAAATGGTTCGAT-GTGGCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAA 495

Qy     360 CAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGTC 407
           ||||| | |||||||| |  ||| || ||   |||||||   | || |
Db     496 CAGTGAACACTCAGGACCCACATGTAAATGCATTTTTCCGTCATTGCC 543
```

RESULT 10
BB222768

```
LOCUS       BB222768                345 bp    mRNA    linear   EST 01-JUL-2000
DEFINITION  BB222768 RIKEN full-length enriched, adult male aorta and vein Mus
            musculus cDNA clone A530079G09 3', mRNA sequence.
ACCESSION   BB222768
VERSION     BB222768.1  GI:8891380
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci,P. and Hayashizaki,Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A530079G09"
                     /clone_lib="RIKEN full-length enriched, adult male aorta
                     and vein"
                     /sex="male"
                     /tissue_type="aorta and vein"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT      105 a     92 c     61 g     87 t
ORIGIN
```

```
  Query Match             23.4%;  Score 117;  DB 10;  Length 345;
  Best Local Similarity   58.9%;  Pred. No. 9.7e-26;
  Matches  201;  Conservative    0;  Mismatches  140;  Indels    0;  Gaps    0;

Qy      65 CTGCATGAAAAGCACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAAGT 124
           || |||  | | || ||||| ||  || | ||   |||  |||  ||||||| |
Db       4 CTCCATCGAGAACATCACCTTTTCTTCTCCTTTCACCACATCACATGTGACATCAATGCC 63

Qy     125 AAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCC 184
           ||||  |   | |  ||||  || |     || ||   | | ||||   | ||||||||
Db      64 AAAGCCAATGCCACAAAGAAAGAGGCTTAGGAGACATCACTGGGCTCTCCATGGCTTCCT 123

Qy     185 CTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTT 244
           ||  ||||| | |      | |   |    ||| |  ||||||| |  |  ||| |||
Db     124 CTGATGAAACATGATCAAATAGCTTCTCAGGAGTACAACATCCCAATAGCAACGACCCTC 183

Qy     245 CCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATT 304
           |||  |    |||  |  ||   | ||| | |   |  |  | ||||  |||    || |||
Db     184 CCTCCCAATTATTTAAGCATTCTAGATCCTACAAGTGCAAAGCATGGTGGAAGTGACATT 243

Qy     305 AAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTG 364
           |||||||| ||||| ||||| || ||  |  ||| | |||||   | | || || ||||||
Db     244 AAATGGGTCGATGGTGGCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACAGTG 303

Qy     365 TATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTG 405
            | |||||||| |  ||| || ||   |||||||   | ||
Db     304 AACACTCAGGACCCACATGTAAATGCATTTTTCCGTCAGTG 344


RESULT 11
BB207661
LOCUS       BB207661                330 bp    mRNA    linear   EST 30-JUN-2000
DEFINITION  BB207661 RIKEN full-length enriched, 0 day neonate thymus Mus
            musculus cDNA clone A430083J17 3', mRNA sequence.
ACCESSION   BB207661
VERSION     BB207661.1  GI:8872614
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 330)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci,P. and Hayashizaki,Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES                Location/Qualifiers
     source             1. .330
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /clone="A430083J17"
                        /clone_lib="RIKEN full-length enriched, 0 day neonate
                        thymus"
                        /tissue_type="thymus"
                        /dev_stage="0 day neonate"
                        /lab_host="DH10B"
                        /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                        prepared and sequenced in Mouse Genome Encyclopedia
                        Project of Genome Exploration Research Group in Riken
                        Genomic Sciences Center and Genome Science Laboratory in
                        RIKEN. Division of Experimental Animal Research in Riken
                        contributed to prepare mouse tissues. 1st strand cDNA was
                        primed with a primer [5'
                        GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                        prepared by using trehalose thermo-activated reverse
                        transcriptase and subsequently enriched for full-length by
                        cap-trapper. cDNA went through one round of normalization
                        to Rot = 10.0 and subtraction to Rot = 459.0. Second
                        strand cDNA was prepared with the primer adapter of
                        sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                        3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                        modified pBluescript KS(+) after bulk excision from Lambda
                        FLC I."
BASE COUNT       96 a     89 c     61 g     84 t
ORIGIN

  Query Match             21.2%;  Score 106;  DB 10;  Length 330;
  Best Local Similarity   57.6%;  Pred. No. 2.8e-22;
  Matches  190;  Conservative    0;  Mismatches  140;  Indels    0;  Gaps    0;

Qy      78 ACCACCTGTTGCTCACATTCTTCCATGTCAGCTGTGACAACTCAAGTAAAGGAAGCACGA 137
           | ||||| ||  || | || | |||  |||  ||||||| |      | ||       | |
Db       1 ATCACCTTTTCTTCTCCTTTTACCACATCACATGTGACATCGATGCCAGAGCCGATCCCA 60

Qy     138 AGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACG 197
           ||||  || |     || ||   | | ||||   | ||||||| ||  ||||| | |
Db      61 GAAAGACAGAGGCTTCGGAGACATCACTGGGCTCTGCATGGCTTCCTCTGATGAAACATG 120

Qy     198 GAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATC 257
           | |   |    ||| |  |||||||     |  ||| ||| ||| |     ||
Db     121 ATCAAATAGCTTCTCAGGAGTACAACATCCCAAAACCAACGACCCTGCCTCCTAATCATT 180

Qy     258 TTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATG 317
           |  ||   ||||| | |   |     | ||||  |||    || ||||||||||| ||||
Db     181 TAAGCATTCAAGATCCTACAAGTCCAAAGCATGGTGGAAGTGACATTAAATGGGTCGATG 240

Qy     318 GAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATC 377
           | | ||| || ||  | ||| | |||||   | | || || |||||| | |||||||| |
Db     241 GTGCCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACAGTGAACACTCAGGACC 300

Qy     378 AGCATTTACATAATTTTTTCCAGTACTGTC 407
           |||  ||  ||   |||||||   | || |
Db     301 CACATGTAAATGCATTTTTCCGTCAGTGCC 330


RESULT 12
BB225426
LOCUS       BB225426                285 bp    mRNA    linear   EST 01-JUL-2000
DEFINITION  BB225426 RIKEN full-length enriched, adult male aorta and vein Mus
            musculus cDNA clone A530092E21 3', mRNA sequence.
ACCESSION   BB225426
```

VERSION BB225426.1 GI:8894037
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 285)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M. , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y. , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y. , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I. , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino ,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .285
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A530092E21"
/clone_lib="RIKEN full-length enriched, adult male aorta and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 88 a 68 c 60 g 69 t
ORIGIN

Query Match 20.0%; Score 100.2; DB 10; Length 285;
Best Local Similarity 59.8%; Pred. No. 1.8e-20;
Matches 168; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```
Qy    125 AAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCC 184
          ||||      | || ||||  || |     || ||   ||| |||||  | |||||||
Db      4 AAAGCCCATGCCAATAAGAAAGAGGCTTCGGAGACATCAGTGGGCTATGCATGGCTTCAT 63

Qy    185 CTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTT 244
          ||  ||||| | |       | |   |    ||| |  |||||||  |  |  ||| |||
Db     64 CTGTTGAAACATGATCAAATAGCTTCTCAGGAGTACAACATCCCAATAGCAACGACCCTG 123

Qy    245 CCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATT 304
          ||| |    ||| |  ||   ||||| | |   |  |  | ||||  |||    || |||
Db    124 CCTCCCAATTATTTAAGCATTCAAGATCCTACAAGTGCAAAGCATGGTGGAAGCGACATT 183

Qy    305 AAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTG 364
          |||||||| ||||| ||||| || ||  | ||| | |||||   | | || || ||||||
Db    184 AAATGGGTCGATGGTGGCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACAGTG 243

Qy    365 TATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTG 405
          | |||||||| |  ||| || ||   |||||||   | ||
Db    244 AACACTCAGGACCCACATGTAAATGCATTTTTCCGTCAGTG 284
```

RESULT 13
BB177025
LOCUS BB177025 274 bp mRNA linear EST 29-JUN-2000
DEFINITION BB177025 RIKEN full-length enriched, adult male hypothalamus Mus musculus cDNA clone A230064N07 3', mRNA sequence.
ACCESSION BB177025
VERSION BB177025.1 GI:8836108
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 274)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M. , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y. , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y. , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I. , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino ,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .274
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A230064N07"
                     /clone_lib="RIKEN full-length enriched, adult male
                     hypothalamus"
                     /sex="male"
                     /tissue_type="hypothalamus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT       83 a     66 c     51 g     74 t
ORIGIN

  Query Match             18.9%;  Score 94.4;  DB 10;  Length 274;
  Best Local Similarity   59.8%;  Pred. No. 1.2e-18;
  Matches  158;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy     135 CGAAGAAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAG 194
           | |  ||||  ||     | || ||   | | |||||  | ||||||||  |  |||||
Db       3 CCACAAAGAAAGAGCCTTTGGAGACATCACTGGGCTATGCATGGCTTCCTTTGTTGAAAC 62

Qy     195 ACGGAAGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCT 254
           | |      | |   |    ||| |  ||||||  |  |  ||| ||| ||| |
Db      63 ATGTTCAAATAGCTTCTCAGGAGTACAACATCCCAATAGCAACGACCCTCCCTCCTAATA 122

Qy     255 ATCTTGGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAG 314
           || |  ||   ||||| | |   |  |  | ||||  |||   || ||||||||||| |
Db     123 ATTTAAGCATTCAAGATCCTACAAGTGCAAAGCATGGTGGAAGTGACATTAAATGGGTCG 182

Qy     315 ATGGAGGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGG 374
           |||| | ||| || ||  | ||| |||||||   | | || || |||||| | |||||||
Db     183 ATGGTGCCAAACCGCTTTTCAAAGTTTCCACATTTGTTGTATCAACAGTGAACACTCAGG 242

Qy     375 ATCAGCATTTACATAATTTTTTCC 398
           | |  ||| || ||   |||||||
Db     243 ACCCACATGTAAATGCATTTTTCC 266
```

RESULT 14
BB219663

```
LOCUS       BB219663                314 bp    mRNA    linear   EST 30-JUN-2000
DEFINITION  BB219663 RIKEN full-length enriched, adult male aorta and vein Mus
            musculus cDNA clone A530056C23 3', mRNA sequence.
ACCESSION   BB219663
VERSION     BB219663.1  GI:8884616
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 314)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci,P. and Hayashizaki,Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .314
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A530056C23"
                     /clone_lib="RIKEN full-length enriched, adult male aorta
                     and vein"
                     /sex="male"
                     /tissue_type="aorta and vein"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
BASE COUNT       85 a     92 c     51 g     86 t
ORIGIN

  Query Match             18.1%;  Score 90.4;  DB 10;  Length 314;
  Best Local Similarity   58.6%;  Pred. No. 2.2e-17;
```

```
  Matches  157;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy     140 AAGAGGGATGTCGTTGAAACCCAAGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGA 199
           ||||  || | | | || ||   ||| ||||   | || ||||| ||  || || | |
Db      47 AAGAAAGAGGCCTTGGAGACATCAGTGGGCTCTCCATGCCTTCCTCTGATGCAACATGAT 106

Qy     200 AGGGTGGTGACAAGCGAGCAGCACATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTT 259
             | |   |    ||| |  || ||||     |  ||| ||| ||| |    ||| |
Db     107 CAAATAGCTTCTCTGGAGTACTACTTCCCCACAGCAACGACCCTCCCTCCTAATTATTTA 166

Qy     260 GGCTACCAAGAGCTTGGGATGGGCAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGA 319
            ||   ||||| | |   |     | || |  |||    || ||||||||||| |||||
Db     167 AGCATTCAAGATCCTACAAGTCCAAAGCTTCGTGGAAGTGACATTAAATGGGTCGATGGT 226

Qy     320 GGCAAGCCACTGCTGAAAATTTCCACTCATCTGGTTTCTACAGTGTATACTCAGGATCAG 379
           | ||| || ||  | ||| | |||||   | | || || |||||| | |||||||| |
Db     227 GCCAAACCGCTTTTCAAAGTGTCCACATTTGTTGTATCAACAGTGAACACTCAGGACCCA 286

Qy     380 CATTTACATAATTTTTTCCAGTACTGTC 407
           ||| || ||   |||||||   | || |
Db     287 CATGTAAATGCATTTTTCCTTCAGTGCC 314


RESULT 15
BB538596
LOCUS       BB538596                 249 bp    mRNA    linear   EST 31-JUL-2000
DEFINITION  BB538596 RIKEN full-length enriched, 0 day neonate eyeball Mus
            musculus cDNA clone E130012K13 3', mRNA sequence.
ACCESSION   BB538596
VERSION     BB538596.1  GI:9594096
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 249)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .249
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="E130012K13"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     eyeball"
                     /tissue_type="eyeball"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT       72 a     65 c     46 g     66 t
ORIGIN

  Query Match             18.0%;  Score 89.8;  DB 10;  Length 249;
  Best Local Similarity   60.4%;  Pred. No. 3.1e-17;
  Matches  148;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy     163 AGTTGGCTACTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCA 222
           ||| | | |  | |||||||| ||  ||||| | |      | |   |    ||| |  |
Db       5 AGTGGCCCATGCATGGCTTCCTCTGATGAAACATGATCAAATAGCTTCTCAGGAGTACAA 64

Qy     223 CATCCCGGTCTCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGG 282
           ||||||  |  |  ||| ||| ||| |    ||| |  ||   | ||| | |   |  |
Db      65 CATCCCCATACCAACGACCCTCCCTCCTAATTATTTAAGCATTCTAGATCCTACAAGTGC 124

Qy     283 CAGGCATTATGGTCCGGAAATTAAATGGGTAGATGGAGGCAAGCCACTGCTGAAAATTTC 342
           | ||||  |||    || ||||||||||| ||||| | ||| || ||  | ||| | ||
Db     125 AAAGCATGGTGGAAGTGACATTAAATGGGTCGATGGTGCCAAACCGCTTTTCAAAGTGTC 184

Qy     343 CACTCATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTA 402
           |||   | | || || |||||| | |||||||| |  ||| || ||   |||||||   |
Db     185 CACATTTGTTGTATCAACAGTGAACACTCAGGACCCACATGTAAATGCATTTTTCCGTCA 244

Qy     403 CTGTC 407
           || |
Db     245 GTGCC 249


Search completed: February  7, 2003, 08:49:31
Job time : 1016.32 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 996.663 Seconds
(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaaagaagtcct.....gaggaggcctccatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	499.4	99.7	2469	9	AK090793	AK090793 Homo sapi
	2	499.4	99.7	3899	6	AX058220	AX058220 Sequence
	3	475.4	94.9	2516	9	AB049841	AB049841 Macaca fa
	4	426.4	85.1	7506	6	AX255048	AX255048 Sequence
	5	426.4	85.1	7522	9	AF527605	AF527605 Homo sapi
	6	426.4	85.1	7545	9	AB028981	AB028981 Homo sapi
	7	366.2	73.1	2998	10	BC009134	BC009134 Mus muscu
	8	359.8	71.8	3227	10	RNTRG	X68101 R.norvegicu
	9	274.4	54.8	2610	6	AX058192	AX058192 Sequence
	10	274.4	54.8	3152	9	AK056684	AK056684 Homo sapi
	11	274.4	54.8	6454	6	AX173022	AX173022 Sequence
	12	269.6	53.8	3472	6	AX173120	AX173120 Sequence
	13	269.6	53.8	4391	6	AX173028	AX173028 Sequence
	14	269.6	53.8	4393	6	AX173118	AX173118 Sequence
	15	263.4	52.6	2299	9	AK054649	AK054649 Homo sapi
	16	248	49.5	139887	9	CNS01RGX	AL160233 Human chr
	17	243.4	48.6	2768	9	AB056820	AB056820 Macaca fa
c	18	222.2	44.4	163316	9	AL161420	AL161420 Human DNA
	19	216.4	43.2	2413	9	BC015018	BC015018 Homo sapi
	20	216.4	43.2	2422	9	AK000227	AK000227 Homo sapi
	21	190.6	38.0	192825	2	AC119357	AC119357 Rattus no
	22	179.4	35.8	182640	2	AC126253	AC126253 Mus muscu
c	23	171.4	34.2	10103	2	AC020032	AC020032 Drosophil
c	24	171.4	34.2	159350	3	AC008354	AC008354 Drosophil
c	25	171.4	34.2	257224	3	AE003617	AE003617 Drosophil
	26	168.2	33.6	1476	6	AX399260	AX399260 Sequence
	27	146.2	29.2	153026	9	AL391280	AL391280 Human DNA
c	28	143	28.5	147556	2	AC110359	AC110359 Rattus no
	29	140.2	28.0	162378	2	AC105541	AC105541 Rattus no
c	30	138.6	27.7	168684	2	AC023985	AC023985 Homo sapi
c	31	138.6	27.7	171811	9	AC011739	AC011739 Homo sapi
	32	138.2	27.6	181127	10	AL672038	AL672038 Mouse DNA
c	33	131.6	26.3	174041	2	AL772340	AL772340 Danio rer
	34	116.4	23.2	2149	6	AX173233	AX173233 Sequence
	35	116.4	23.2	2825	9	BC008335	BC008335 Homo sapi
	36	116.4	23.2	6372	6	AX173175	AX173175 Sequence
	37	112.8	22.5	1900	10	BC021414	BC021414 Mus muscu
	38	102.6	20.5	4886	9	AB037816	AB037816 Homo sapi
	39	93.4	18.6	2303	10	BC024917	BC024917 Mus muscu
	40	93.4	18.6	2972	10	BC024823	BC024823 Mus muscu
c	41	88.2	17.6	73515	2	AC017375	AC017375 Drosophil
	42	88.2	17.6	166626	3	AC008318	AC008318 Drosophil
	43	88.2	17.6	301639	3	AE003590	AE003590 Drosophil
	44	83.2	16.6	2014	6	AX174662	AX174662 Sequence
	45	83.2	16.6	4026	6	AX174571	AX174571 Sequence

ALIGNMENTS

RESULT 1
AK090793
LOCUS AK090793 2469 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ33474 fis, clone BRAMY2002369, highly similar to Trg protein.
ACCESSION AK090793
VERSION AK090793.1 GI:21749019
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens amygdala cDNA to mRNA, clone_lib:BRAMY2 clone:BRAMY2002369.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,

```
            Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
            Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2469)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .2469
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BRAMY2002369"
                     /tissue_type="amygdala"
                     /clone_lib="BRAMY2"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      666 a    561 c    556 g    686 t
ORIGIN

  Query Match             99.7%;  Score 499.4;  DB 9;  Length 2469;
  Best Local Similarity   99.8%;  Pred. No. 2.7e-122;
  Matches  500;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 1700

Qy     61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1701 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 1760

Qy    121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 1820

Qy    181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1821 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 1880

Qy    241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1881 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 1940

Qy    301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1941 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2000

Qy    361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2001 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2060

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2061 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2120

Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   2121 CGAGGAGGCCTCCATGATGGA 2141
```

```
RESULT 2
AX058220
LOCUS       AX058220               3899 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 90 from Patent WO0077040.
ACCESSION   AX058220
VERSION     AX058220.1  GI:12310721
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3899)
  AUTHORS   Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Yang,J., Reddy,R. and Lu,D.A.
  TITLE     Human intracellular signaling molecules
  JOURNAL   Patent: WO 0077040-A 90 21-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3899
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1078 a    929 c    949 g    943 t
ORIGIN

  Query Match             99.7%;  Score 499.4;  DB 6;  Length 3899;
  Best Local Similarity   99.8%;  Pred. No. 2.7e-122;
  Matches  500;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1922 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 1981

Qy     61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1982 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 2041

Qy    121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2042 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2101

Qy    181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2102 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2161

Qy    241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2162 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2221

Qy    301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2222 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2281

Qy    361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2282 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2341

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2342 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2401

Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   2402 CGAGGAGGCCTCCATGATGGA 2422
```

```
RESULT 3
AB049841
LOCUS       AB049841               2516 bp    mRNA    linear   PRI 14-OCT-2000
DEFINITION  Macaca fascicularis brain cDNA, clone:QnpA-17096.
ACCESSION   AB049841
VERSION     AB049841.1  GI:10801619
```

KEYWORDS fis (full insert sequence).
SOURCE Macaca fascicularis adult male brain parietal lobe cDNA to mRNA, clone_lib:macaque brain cDNA library QnpA clone:QnpA-17096.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1 (sites)
AUTHORS Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M., Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
TITLE Isolation of full-length cDNA clones from macaque brain cDNA libraries
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2516)
AUTHORS Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2000) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/, Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGTGTG)
R. Site2: DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb.The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by Sugano et al.(University of Tokyo, Institute of Medical Science). Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAAGCTGCG];
3' end primer [CGACCTGCAGCTCGAGCACA]).
FEATURES Location/Qualifiers
source 1. .2516
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="QnpA-17096"
/sex="male"
/tissue_type="brain parietal lobe"
/clone_lib="macaque brain cDNA library QnpA"
/dev_stage="adult"
CDS 1242. .>2516
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB16727.1"
/db_xref="GI:10801620"
/translation="MMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVC
LTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRPL
IYKFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSF
VRTHLQVIISVSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKR
IRTVLMATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSE
AAMCYVHVTALVAEYLTRKGMFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVL
MELLEQCADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVM
HSGRRLLGTYFRVAFFGQVSLLSVLQCPKSFKE"
BASE COUNT 657 a 578 c 572 g 709 t
ORIGIN

Query Match 94.9%; Score 475.4; DB 9; Length 2516;
Best Local Similarity 96.8%; Pred. No. 6.6e-116;
Matches 485; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    1700 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCCGTCAGCCA 1759

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    1760 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 1819

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||| || ||||||||||||||||||| ||||||||
Db    1820 CAACAACTGTGCCAACAGTGACCGGCTCATCAAGCACACCAGCTTCTCCTCCGATGTGAA 1879

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Db    1880 GGACTTGACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAACACGA 1939

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1940 GAATGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 1999

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||| |||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||
Db    2000 GCCTGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCACGTCAAAAACGGCGA 2059

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          |||||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||
Db    2060 TCTCTCAGAGGCAGCAATGTGCTATGTCCATGTAACAGCCCTGGTGGCAGAATATCTCAC 2119

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2120 CCGGAAAGGCATGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2179

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2180 CGAGGAGGCCTCCATGATGGA 2200
```

RESULT 4
AX255048
LOCUS AX255048 7506 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 7 from Patent WO0170808.
ACCESSION AX255048
VERSION AX255048.1 GI:16074541
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7506)
AUTHORS Rastelli,L.K. and Gerritsen,M.
TITLE Angiogenesis-associated proteins, and nucleic acids encoding the same
JOURNAL Patent: WO 0170808-A 7 27-SEP-2001;
Curagen Corporation (US) ; GENENTECH, INC. (US)
FEATURES Location/Qualifiers
source 1. .7506
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2147 a 1689 c 1764 g 1906 t
ORIGIN

Query Match 85.1%; Score 426.4; DB 6; Length 7506;
Best Local Similarity 99.8%; Pred. No. 8.1e-103;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4572 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 4631

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    4632 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 4691

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4692 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 4751

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4752 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 4811
```

```
Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4812 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 4871

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4872 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 4931

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4932 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 4991

Qy     421 ACGGAAAG 428
           ||||||||
Db    4992 ACGGAAAG 4999
```

```
RESULT 5
AF527605
LOCUS       AF527605               7522 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens zizimin1 mRNA, complete cds.
ACCESSION   AF527605
VERSION     AF527605.1  GI:22038158
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7522)
  AUTHORS   Meller,N., Irani-Tehrani,M., Kiosses,W.B., Del Pozo,M.A. and
            Schwartz,M.A.
  TITLE     Zizimin1, a novel Cdc42 activator, reveals new guanine nucleotide
            exchange-exchange factor domain for rho proteins
  JOURNAL   Nat. Cell Biol. (2002) In press
REFERENCE   2  (bases 1 to 7522)
  AUTHORS   Meller,N. and Schwartz,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2002) Cell Biology, The Scripps Research
            Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .7522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
     CDS             56. .6265
                     /note="guanine nucleotide exchange factor; Cdc42
                     activator"
                     /codon_start=1
                     /product="zizimin1"
                     /protein_id="AAM90306.1"
                     /db_xref="GI:22038159"
                     /translation="MSQPPLLPASAETRKFTRALSKPGTAAELRQSVSEVVRGSVLLA
                     KPKLIEPLDYENVIVQKKTQILNDCLREMLLFPYDDFQTAILRRQGRYICSTVPAKAE
                     EEAQSLFVTECIKTYNSDWHLVNYKYEDYSGEFRQLPNKVVKLDKLPVHVYEVDEEVD
                     KDEDAASLGSQKGGITKHGWLYKGNMNSAISVTMRSFKRRFFHLIQLGDGSYNLNFYK
                     DEKISKEPKGSIFLDSCMGVVQNNKVRRFAFELKMQDKSSYLLAADSEVEMEEWITIL
                     NKILQLNFEAAMQEKRNGDSHEDDEQSKLEGSGSGLDSYLPELAKSAREAEIKLKSES
                     RVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGKRILVKCNDLSFNLQCCVAENEEGPTT
                     NVEPFFVTLSLFDIKYNRKISADFHVDLNHFSVRQMLATTSPALMNGSGQSPSVLKGI
                     LHEAAMQYPKQGIFSVTCPHPDIFLVARIEKVLQGSITHCAEPYMKSSDSSKVAQKVL
                     KNAKQACQRLGQYRMPFAWAARTLFKDASGNLDKNARFSAIYRQDSNKLSNDDMLKLL
                     ADFRKPEKMAKLPVILGNLDITIDNVSSDFPNYVNSSYIPTKQFETCSKTPITFEVEE
                     FVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQKSFAKARNIAICIEFKDSDEEDSQPLK
                     CIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEIKIELPTQLHEKHHLLLTFFHVSCDNS
                     SKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHIPVSANLPSGYLGYQELGMGRHYG
                     PEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYCQKTESGAQALGNELVKYLKSL
                     HAMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVIIHVVAQCHEEGLESHLRSY
                     VKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNKLLKYSWFFFDVLIKSMA
                     QHLIENSKVKLLRNQRFPASYHHAVETVVNMLMPHITQKFRDNPEASKNANHSLAVFI
                     KRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNHEHYIPLNLPMPFG
                     KGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLREVGTALQEFREVRLIAISVLKNLLIK
                     HSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMTVKDESLALP
                     AVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGSLISTDSG
                     NSLPERNSEKSNSLDKHQQSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALF
                     TYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARTGMMHARLQQLGSLDNSLTFNHS
                     YGHSDADVLHQSLLEANIATEVCLTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVY
                     LCFLQKHQSETALKNVFTALRSLIYKFPSTFYEGRADMCAALCYEILKCCNSKLSSIR
                     TEASQLLYFLMRNNFDYTGKKSFVRTHLQVIISVSQLIADVVGIGGTRFQQSLSIINN
                     CANSDRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHENDPEMLVDLQYSLAKSYAST
                     PELRKTWLDSMARIHVKNGDLSEAAMCYVHVTALVAEYLTRKEAVQWEPPLLPHSHSA
                     CLRRSRGGVFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVLMELLEQCADGLW
                     KAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVMHSGRRLLGTYFR
                     VAFFGQAAQYQFTDSETDVEGFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGS
                     ENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEM
                     PFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKK
                     VAELRQLCSSAEVDMIKLQLKLQGSVSVQVNAGPLAYARAFLDDTNTKRYPDNKVKLL
                     KEVFRQFVEACGQALAVNERLIKEDQLEYQEEMKANYREMAKELSEIMHEQLG"
BASE COUNT     2146 a    1694 c    1774 g    1908 t
ORIGIN

  Query Match           85.1%;  Score 426.4;  DB 9;  Length 7522;
  Best Local Similarity  99.8%;  Pred. No. 8.1e-103;
  Matches  427;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4585 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 4644

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    4645 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 4704

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4705 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 4764

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4765 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 4824

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4825 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 4884

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4885 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 4944

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4945 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 5004

Qy     421 ACGGAAAG 428
           ||||||||
Db    5005 ACGGAAAG 5012
```

```
RESULT 6
AB028981
LOCUS       AB028981               7545 bp    mRNA    linear   PRI 10-MAY-2002
DEFINITION  Homo sapiens mRNA for KIAA1058 protein, partial cds.
ACCESSION   AB028981
VERSION     AB028981.2  GI:20521745
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh12146s1.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
```

```
  MEDLINE   99397452
   PUBMED   10470851
REFERENCE   2  (bases 1 to 7545)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
COMMENT     On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES             Location/Qualifiers
     source          1. .7545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="hh12146s1"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
                     /note="This sequence was obtained bysubcloning of the DNA
                     fragments derived from two cDNA clones (1 - 1846 was
                     derived from ef00877 and 1847 - 7545 was derived from
                     hh12146)."
     gene            1. .7545
                     /gene="KIAA1058"
     CDS             <1. .6288
                     /gene="KIAA1058"
                     /codon_start=1
                     /product="KIAA1058 protein"
                     /protein_id="BAA83010.2"
                     /db_xref="GI:20521746"
                     /translation="LACLCLARPGTRSCSSCRLSDLRFQPRMQADKCRTSSRSVKKEL
                     VIESPLQYKDAAQGEVEAESPGPVPAKPKLIEPLDYENVIVQKKTQILNDCLREMLLF
                     PYDDFQTAILRRQGRYICSTVPAKAEEEAQSLFVTECIKTYNSDWHLVNYKYEDYSGE
                     FRQLPNKVVKLDKLPVHVYEVDEEVDKDEDAASLGSQKGGITKHGWLYKGNMNSAISV
                     TMRSFKRRFFHLIQLGDGSYNLNFYKDEKISKEPKGSIFLDSCMGVVQNNKVRRFAFE
                     LKMQDKSSYLLAADSEVEMEEWITILNKILQLNFEAAMQEKRNGDSHEDDEQSKLEGS
                     GSGLDSYLPELAKSAREAEIKLKSESRVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGK
                     RILVKCNDLSFNLQCCVAENEEGPTTNVEPFFVTLSLFDIKYNRKISADFHVDLNHFS
                     VRQMLATTSPALMNGSGQSPSVLKGILHEAAMQYPKQGIFSVTCPHPDIFLVARIEKV
                     LQGSITHCAEPYMKSSDSSKVAQKVLKNAKQACQRLGQYRMPFAWAARTLFKDASGNL
                     DKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSDFPN
                     YVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQK
                     SFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEI
                     KIELPTQLHEKHHLLLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSE
                     QHIPVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHN
                     FFQYCQKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLTRATQEEVA
                     VNVTRVIIHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKP
                     SADFLTSNKLLKYSWFFFDVLIKSMAQHLIENSKVKLLRNQRFPASYHHAVETVVNML
                     MPHITQKFRDNPEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLF
                     EYKFEFLRVVCNHEHYIPLNLPMPFGKGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLR
                     EVGTALQEFREVRLIAISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQ
                     RINVRDVSPFPVNAGMTVKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASP
                     YTTSTPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSSTLGNSVVRCDK
                     LDQSEIKSLLMCFLYILKSMSDDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRY
                     IARTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSL
                     FTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRSLIYKFPSTFY
                     EGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSFVRTHLQVII
                     SVSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKRIRTVLMATA
                     QMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSEAAMCYVHVT
                     ALVAEYLTRKEAVQWEPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPNIDEEASMME
                     DVGMQDVHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLY
                     DTLHRAYSKVTEVMHSGRRLLGTYFRVAFFGQAAQYQFTDSETDVEGFFEDEDGKEYI
                     YKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPF
                     FDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVK
                     KRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLCSSAEVDMIKLQLKLQGSVSVQVNA
                     GPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGQALAVNERLIKEDQLEYQEE
                     MKANYREMAKELSEIMHEQLG"
BASE COUNT     2175 a   1687 c   1753 g   1930 t
ORIGIN
```

Query Match 85.1%; Score 426.4; DB 9; Length 7545;
Best Local Similarity 99.8%; Pred. No. 8.1e-103;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4608 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 4667

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    4668 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 4727

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4728 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 4787

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4788 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 4847

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4848 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 4907

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4908 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 4967

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4968 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 5027

Qy     421 ACGGAAAG 428
          ||||||||
Db    5028 ACGGAAAG 5035
```

```
RESULT 7
BC009134
LOCUS       BC009134                2998 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to hypothetical protein FLJ20220, clone
            MGC:11827 IMAGE:3596515, mRNA, complete cds.
ACCESSION   BC009134
VERSION     BC009134.1  GI:14318664
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2998)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 16 Row: i Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
```

```
          analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .2998
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:11827 IMAGE:3596515"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             381. .1634
                     /codon_start=1
                     /product="Similar to hypothetical protein FLJ20220"
                     /protein_id="AAH09134.1"
                     /db_xref="GI:14318665"
                     /translation="MFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVLMELLEQ
                     CADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVMHSGRRL
                     LGTYFRVAFFGQGFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQD
                     SGKVNPKDLDSKFAYIQVTHVTPFFDEKELQERRTEFERCHNIRRFMFEMPFTQTGKR
                     QGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLC
                     SSAEVDMIKLQLKLQGSVSVQVNAGPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFV
                     EACGQALAVNERLIKEDQLEYQEEMKANYREMAKELSDIMREQICPLEEKTSVLPNSL
                     HIFNAISGTPTSTVVQGLTSSSSVV"
BASE COUNT      835 a    655 c    768 g    740 t
ORIGIN

  Query Match             73.1%;  Score 366.2;  DB 10;  Length 2998;
  Best Local Similarity   88.2%;  Pred. No. 8.5e-87;
  Matches  398;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy      51 CTGTCAGCCAGCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCC 110
           |||||||||| ||||| ||||| || |||||||||| | |||||||||||||||||||
Db       1 CTGTCAGCCAACTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCT 60

Qy     111 TGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCT 170
           |||| |||||||||||||||||||||| |||||| | || |||||||||||||| ||||
Db      61 TGTCTATCATCAACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCT 120

Qy     171 CTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGA 230
           |||||||||| || || || || ||||| ||||| || || |||||||| ||||||||||
Db     121 CTGATGTGAAAGATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGA 180

Qy     231 AGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCT 290
           ||||||| |||||||||| ||||||||||||||||||||||||||||||| || ||||
Db     181 AGGAGCACGAGAACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCT 240

Qy     291 ATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCA 350
           | |||||||| || |||||||||||||||||||| ||||| ||||| ||||| || || |
Db     241 ACGCCAGCACCCCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTA 300

Qy     351 AAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAG 410
           ||||||| || |||||||||| |||||||||||||||||||| |||||| | |||||||
Db     301 AAAATGGGGACCTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAG 360

Qy     411 AATATCTCACACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCC 470
           |||||||||||||||||||| |||| ||||| || ||||| |||||||||| || || |
Db     361 AATATCTCACACGGAAAGGCATGTTCAGACAGGGGTGCACAGCCTTCAGGGTTATCACAC 420

Qy     471 CAAACATCGACGAGGAGGCCTCCATGATGGA 501
           |||||||||| || ||||| ||||||||||
Db     421 CAAACATCGATGAAGAGGCTTCCATGATGGA 451


RESULT 8
RNTRG
LOCUS       RNTRG                   3227 bp    mRNA    linear   ROD 24-FEB-1999
DEFINITION  R.norvegicus trg mRNA.
ACCESSION   X68101
VERSION     X68101.1  GI:550419
KEYWORDS    trg gene.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3227)
  AUTHORS   Pianese,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1992) L. Pianese, Dipartimento di Biologia,
            Patologia Cellulare e Molecolare, Via Pansini, Naples, ITALY
REFERENCE   2  (bases 1 to 3227)
  AUTHORS   Pianese,L., Porcellini,A., Avvedimento,V.E., D'Esposti,F.,
            Feliciello,A., Monticelli,A., Musti,A.M., Tortora,G., Varrone,S.
            and Cocozza,S.
  TITLE     A novel thyroid transcript negatively regulated by tsh
  JOURNAL   Mol. Biol. 13, 75-83 (1994)
FEATURES             Location/Qualifiers
     source          1. .3227
                     /organism="Rattus norvegicus"
                     /strain="Fischer"
                     /db_xref="taxon:10116"
                     /cell_line="FRTL-"
                     /clone_lib="lambdaGT"
     gene            1. .2218
                     /gene="trg"
     CDS             <1. .2218
                     /gene="trg"
                     /codon_start=2
                     /protein_id="CAA48220.1"
                     /db_xref="GI:550420"
                     /db_xref="SPTREMBL:Q63603"
                     /translation="KLSRGHSPLMKKVFDVYLCFLQKHQSEMALKNVFTALRSLIYKF
                     PSTFYEGRADMCASLCYEVLKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSFVRTH
                     LQVIISLSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKRIRTV
                     LMATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSEAAMC
                     YVHVTALVAEYLTRKEADLALQREPPVFPYSHTSCQRKSRGGMFRQGCTAFRVITPNI
                     DEEASMMEDVGMQDVHFNEDVLMELLEQCADGLWKAERLRAGLLTSINSSSPSMKSGG
                     TLETTHLYDTLHRPYSKVTEVITRAAGSWDLLPGGLFGQGFFEDEDGKEYIYKEPKLT
                     PLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKFAYIQVTHVTPFFDEKELQ
                     ERKTEFERCHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMY
                     QHHTDLNPIEVAIDEMSKKVAELHQLCSSAEVDMIKLQLKLQGSVSVQVNAGPLAYAR
                     AFLDDTNTKRYPDNKVKLLKEVFRQFVEACGQALAVNERLIKEDQLEYQEEMKANYRE
                     IRKELSDIIVPRICPGEDKRATKFPAHLQRHQRDTNKHSGSRVDQFILSCVTLPHEPH
                     VGTCFVMCKLRTTFRANHWFCQAQEEAMGNGREKEPWTVIFNSRFYRSWGKVHIFF"
BASE COUNT      905 a    738 c    777 g    807 t
ORIGIN

  Query Match             71.8%;  Score 359.8;  DB 10;  Length 3227;
  Best Local Similarity   90.2%;  Pred. No. 4.3e-85;
  Matches  385;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||||| || |||||||| |||||||||||||| || || |||||||| ||| ||||||||
Db     275 TACACAGGGAAGAAGTCTTTTGTCCGGACACACTTACAGGTCATCATCTCTCTCAGCCAG 334

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| |||||||| |||||||||| | |||||||||||||||||| |||| ||||||
Db     335 CTGATCGCAGACGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 394

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||| || |||||||| ||||||||||||||||||||||||||||||||
Db     395 AACAACTGTGCCAATAGCGACCGGCTCATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 454

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || ||||| ||||| ||||| || || |||||||| |||||||||||||| ||||||
Db     455 GATTTGACCAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAACATGAG 514

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||||||||||||||||||||| |||||||||||||||| ||||| ||||||||
Db     515 AACGACCCAGAGATGCTGGTGGACCTGCAGTACAGCCTGGCCAAGTCCTACGCCAGCACC 574

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||||||| ||||| |||||||||| |||||||||| |||||||| |||
Db     575 CCTGAGCTCAGGAAGACATGGCTGGACAGCATGGCAAGGATCCATGTTAAAAATGGGGAT 634
```

```
Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           |||||||||||||||||||||||||||||||| ||||| | || || ||||||||||||
Db     635 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTGACAGCTTTGGTAGCGGAATATCTCACA 694

Qy     422 CGGAAAG 428
           |||||||
Db     695 CGGAAAG 701
```

```
RESULT 9
AX058192
LOCUS       AX058192                2610 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 62 from Patent WO0077040.
ACCESSION   AX058192
VERSION     AX058192.1  GI:12310693
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2610)
  AUTHORS   Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Yang,J., Reddy,R. and Lu,D.A.
  TITLE     Human intracellular signaling molecules
  JOURNAL   Patent: WO 0077040-A 62 21-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2610
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      878 a    467 c    529 g    736 t
ORIGIN
```

```
  Query Match             54.8%;  Score 274.4;  DB 6;  Length 2610;
  Best Local Similarity   71.8%;  Pred. No. 2.5e-62;
  Matches  359;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db     591 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 650

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db     651 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 710

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           || || | ||| || |||||| | | ||| |||   ||   ||| || | || || ||
Db     711 AATAATTTTGCAAATAGTGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAA 770

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||||| ||||| || || |||| || || |||||||| |||| ||||||||||||||||
Db     771 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 830

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db     831 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 890

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db     891 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 950

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
            | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db     951 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 1010

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           || |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db    1011 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 1070

Qy     482 GAGGAGGCCTCCATGATGGA 501
           || || |   | ||||  ||
Db    1071 GAAGAAGGAGCAATGAAAGA 1090
```

```
RESULT 10
AK056684
LOCUS       AK056684                3152 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ32122 fis, clone PEBLM1000144, moderately
            similar to Trg.
ACCESSION   AK056684
VERSION     AK056684.1  GI:16552156
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens peripheral blood mononuclear cells (HPBMC5939) cDNA to
            mRNA, clone_lib:PEBLM1 clone:PEBLM1000144.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
            Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
            Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
            Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
            Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
            Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
            Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
            Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3152)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .3152
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PEBLM1000144"
                     /cell_type="peripheral blood mononuclear cells
                     (HPBMC5939)"
                     /clone_lib="PEBLM1"
                     /note="cloning vector: pME18SFL3~primary culture,
                     peripheral blood mononuclear cells"
     CDS             154. .2799
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB71253.1"
                     /db_xref="GI:16552157"
                     /translation="MPNSASRDEFPCGFTSPANRGSLSTDKDTAYGSFQNGHGIKRED
                     SRGSLIPEGATGFPDQGNTGENTRQSSTRSSVSQYNRLDQYEIRSLLMCYLYIVKMIS
                     EDTLLTYWNKVSPQELINILILLEVCLFHFRYMGKRNIARVHDAWLSKHFGIDRKSQT
                     MPALRNRSGVMQARLQHLSSLESSFTLNHSSTTTEADIFHQALLEGNTATEVSLTVLD
                     TISFFTQCFKTQLLNNDGHNPLMKKVFDIHLAFLKNGQSEVSLKHVFASLRAFISKFP
                     SAFFKGRVNMCAAFCYEVLKCCTSKISSTRNEASALLYLLMRNNFEYTKRKTFLRTHL
                     QIIIAVSQLIADVALSGGSRFQESLFIINNFANSDRPMKATAFPAEVKDLTKRIRTVL
                     MATAQMKEHEKDPEMLIDLQYSLAKSYASTPELRKTWLDSMAKIHVKNGDFSEAAMCY
                     VHVAALVAEFLHRKKLFPNGCSAFKKITPNIDEEGAMKEDAGMMDVHYSEEVLLELLE
                     QCVDGLWKAERYEIISEISKLIVPIYEKRREFEKLTQVYRTLHGAYTKILEVMHTKKR
                     LLGTFFRVAFYGQSFFEEEDGKEYIYKEPKLTGLSEISLRLVKLYGEKFGTENVKIIQ
                     DSDKVNAKELDPKYAHIQVTYVKPYFDDKELTERKTEFERNHNISRFVFEAPYTLSGK
                     KQGCIEEQCKRRTILTTSNSFPYVKKRIPINCEQQINLKPIDVATDEIKDKTAELQKL
                     CSSTDVDMIQLQLKLQGCVSVQVNAGPLAYARAFLNDSQASKYPPKKVSELEDMFRKF
                     IQACSIALELNERLIKEDQVEYHEGLKSNFRDMVKELSDIIHEQILQEDTMHSPWMSN
                     TLHVFCAISGTSSDRGYGSPRYAEV"
BASE COUNT      1058 a    576 c    638 g    880 t
```

```
ORIGIN

  Query Match             54.8%;  Score 274.4;  DB 9;  Length 3152;
  Best Local Similarity   71.8%;  Pred. No. 2.5e-62;
  Matches  359;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||  ||   ||  |||  |||||  |  ||||||||  |  ||   |  ||  ||   ||||  |||||
Db    1120 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 1179

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||||||  ||  ||  |     |  ||  |  |  |||||  |||  ||||  |  |  |||  |||
Db    1180 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 1239

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           ||  ||  |  |||  ||  ||||||  |  |  |||  |||    ||    |||   ||  |  ||  ||  ||
Db    1240 AATAATTTTGCAAATAGTGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAA 1299

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           |||||  |||||  ||  ||  |||||  ||  ||  ||||||||  |||||  ||||||||||||||||
Db    1300 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 1359

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           ||  |||||  ||  |||||   |  ||  ||||||||  |||  |  |||||  ||||||||  |||||
Db    1360 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 1419

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           ||  ||||||||||||  ||  |||||  ||  ||||||||||  |||  |||||  ||||||||  |||
Db    1420 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 1479

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
             |  ||||||||  ||  |||||  ||||||||  |||  ||||  |||||  |||||  |  |||
Db    1480 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 1539

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           ||  |||    |  |||    |  |||||  |  ||  ||||  |   |||||  ||  ||  ||  ||
Db    1540 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 1599

Qy     482 GAGGAGGCCTCCATGATGGA 501
           ||  ||  |   |  ||||  ||
Db    1600 GAAGAAGGAGCAATGAAAGA 1619


RESULT 11
AX173022
LOCUS       AX173022                6454 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 1 from Patent WO0142294.
ACCESSION   AX173022
VERSION     AX173022.1  GI:14597975
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6454)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 1 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6454
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="full length human CLASP-4 cDNA"
     CDS             95. .6121
                     /note="human CLASP-4"
                     /codon_start=1
                     /protein_id="CAC43729.1"
                     /db_xref="GI:14597976"
                     /translation="MFPMEDISISVIGRQRRTVQSTVPEDAEKRAQSLFVKECIKTYS
                     TDWHVVNYKYEDFSGDFRMLPCKSLRPEKIPNHVFEIDEDCEKDEDSSSLCSQKGGVI
                     KQGWLHKANVNSTITVTMKVFKRRYFYLTQLPDGSYILNSYKDEKNSKESKGCIYLDA
                     CIDVVQCPKMRRHAFELKMLDKYSHYLAAETEQEMEEWLITLKKIIQINTDSLVQEKK
                     ETVETAQDDETSSQGKAENIMASLERSMHPELMKYGRETEQLNKLSRGDGRQNLFSFD
                     SEVQRLDFSGIEPDIKPFEEKCNKRFLVNCHDLTFNILGQIGDNAKGPPTNVEPFFIN
                     LALFDVKNNCKISADFHVDLNPPSVREMLWGSSTQLASDGSPKGSSPESYIHGIAESQ
                     LRYIQQGIFSVTNPHPEIFLVARIEKVLQGNITHCAEPYIKNSDPVKTAQKVHRTAKQ
                     VCSRLGQYRMPFAWAARPIFKDTQGSLDLDGRFSPLYKQDSSKLSSEDILKLLSEYKK
                     PEKTKLQIIPGQLNITVECVPVDLSNCITSSYVPLKPFEKNCQNITVEVEEFVPEMTK
                     YCYPFTIYKNHLYVYPLQLKYDSQKTFAKARNIAVCVEFRDSDESDASALKCIYGKPA
                     GSVFTTNAYAVVSHHNQNPEFYDEIKIELPIHLHQKHHLLFTFYHVSCEINTKGTTKK
                     QDTVETPVGFAWVPLLKDGRIITFEQQLPVSANLPPGYLNLNDAESRRQCNVDIKWVD
                     GAKPLLKFKSHLESTIYTQDLHVHKFFHHCQLIQSGSKEVPGELIKYLKCLHAMEIQV
                     MIQFLPVILMQLFRVLTNMTHEDDVPINCTMVLLHIVSKCHEEGLDSYLRSFIKYSFR
                     PEKPSAPQAQLIHETLATTMIAILKQSADFLSINKLLKYSWFFFEIIAKSMATYLLEE
                     NKIKLPRGQRFPETYHHVLHSLLLAIIPHVTIRYAEIPDESRNVNYSLASFLKRCLTL
                     MDRGFIFNLINDYISGFSPKDPKVLAEYKFEFLQTICNHEHYIPLNLPMAFAKPKLQR
                     VQDSNLEYSLSDEYCKHHFLVGLLLRETSIALQDNYEIRYTAISVIKNLLIKHAFDTR
                     YQHKNQQAKIAQLYLPFVGLLLENIQRLAGRDTLYSCAAMPNSASRDEFPCGFTSPAN
                     RGSLSTDKDTAYGSFQNGHGIKREDSRGSLIPEGATGFPDQGNTGENTRQSSTRSSVS
                     QYNRLDQYEIRSLLMCYLYIVKMISEDTLLTYWNKVSPQELINILILLEVCLFHFRYM
                     GKRNIARVHDAWLSKHFGIDRKSQTMPALRNRSGVMQARLQHLSSLESSFTLNHSSTT
                     TEADIFHQALLEGNTATEVSLTVLDTISFFTQCFKTQLLNNDGHNPLMKKVFDIHLAF
                     LKNGQSEVSLKHVFASLRAFISKFPSAFFKGRVNMCAAFCYEVLKCCTSKISSTRNEA
                     SALLYLLMRNNFEYTKRKTFLRTHLQIIIAVSQLIADVALSGGSRFQESLFIINNFAN
                     SDRPMKATAFPAEVKDLTKRIRTVLMATAQMKEHEKDPEMLIDLQYSLAKSYASTPEL
                     RKTWLDSMAKIHVKNGDFSEAAMCYVHVAALVAEFLHRKKLFPNGCSAFKKITPNIDE
                     EGAMKEDAGMMDVHYSEEVLLELLEQCVDGLWKAERYEIISEISKLIVPIYEKRREFE
                     KLTQVYRTLHGAYTKILEVMHTKKRLLGTFFRVAFYGQSFFEEEDGKEYIYKEPKLTG
                     LSEISLRLVKLYGEKFGTENVKIIQDSDKVNAKELDPKYAHIQVTYVKPYFDDKELTE
                     RKTEFERNHNISRFVFEAPYTLSGKKQGCIEEQCKRRTILTTSNSFPYVKKRIPINCE
                     QQINLKPIDGATDEIKDKTAELQKLCSSTDVDMIQLQLKLQGWVSVQVNAGPLAYARA
                     FLNDSQASKYPPKKVSELKDMFRKFIQACSIALELNERLIKEDQVEYHEGLKSNFRDM
                     VKELSDIIHEQILQEDTMHSPWMSNTLHVFCAISGTSSDRGYGSPRYAEV"
BASE COUNT      2119 a    1234 c    1303 g    1798 t
ORIGIN

  Query Match             54.8%;  Score 274.4;  DB 6;  Length 6454;
  Best Local Similarity   71.8%;  Pred. No. 2.6e-62;
  Matches  359;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||  ||   ||  |||  |||||  |  ||||||||  |  ||   |  ||  ||   ||||  |||||
Db    4442 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 4501

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||  ||  ||  |     |  ||  |  |  |||||  |||  ||||  |  |  |||  |||
Db    4502 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 4561

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           ||  ||  |  |||  ||  ||||||  |  |  |||  |||    ||    |||   ||  |  ||  ||  ||
Db    4562 AATAATTTTGCAAATAGTGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAA 4621

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           |||||  |||||  ||  ||  |||||  ||  ||  ||||||||  |||||  ||||||||||||||||
Db    4622 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 4681

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           ||  |||||  ||  |||||   |  ||  ||||||||  |||  |  |||||  ||||||||  |||||
Db    4682 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 4741

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           ||  ||||||||||||  ||  |||||  ||  ||||||||||  |||  |||||  ||||||||  |||
Db    4742 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 4801

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
             |  ||||||||  ||  |||||  ||||||||  |||  ||||  |||||  |||||  |  |||
Db    4802 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 4861

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           ||  |||    |  |||    |  |||||  |  ||  ||||  |   |||||  ||  ||  ||  ||
Db    4862 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 4921

Qy     482 GAGGAGGCCTCCATGATGGA 501
           ||  ||  |   |  ||||  ||
Db    4922 GAAGAAGGAGCAATGAAAGA 4941
```

```
RESULT 12
AX173120
LOCUS       AX173120               3472 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 99 from Patent WO0142294.
ACCESSION   AX173120
VERSION     AX173120.1  GI:14598013
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3472)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 99 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .3472
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="polynucleotide fragment"
BASE COUNT     1153 a    640 c    696 g    983 t
ORIGIN

  Query Match           53.8%;  Score 269.6;  DB 6;  Length 3472;
  Best Local Similarity  71.2%;  Pred. No. 4.8e-61;
  Matches  356;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db    1461 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 1520

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db    1521 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 1580

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           || || | ||| || |||||| | | |||   |   |    |||  || | || || ||
Db    1581 AATAATTTTGCAAATAGTGACAGACCTATGTTGGCAAGAGCCTTTCCCGCAGAAGTCAAA 1640

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||||| ||||| || || ||||| || || |||||||| ||||| ||||||||||||||||
Db    1641 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 1700

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db    1701 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 1760

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db    1761 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 1820

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db    1821 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 1880

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           || |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db    1881 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 1940

Qy     482 GAGGAGGCCTCCATGATGGA 501
           || || |   | ||||  ||
Db    1941 GAAGAAGGAGCAATGAAAGA 1960


RESULT 13
AX173028
LOCUS       AX173028               4391 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 7 from Patent WO0142294.
ACCESSION   AX173028
VERSION     AX173028.1  GI:14597978
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4391)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 7 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .4391
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="human CLASP-4 cDNA"
     CDS             414. .4058
                     /note="human CLASP-4"
                     /codon_start=1
                     /protein_id="CAC43730.1"
                     /db_xref="GI:14597979"
                     /translation="MEIQVLIRFLSVILMQLFWVLPNMIHEDDVPISCPMVLFHIVSK
                     CHEEGLDSYLSSFIKYSFRPGKPSAPQAPLIHETLATMMIALLKQSADFLAINKLLKY
                     SWFFFEIIAKSMATYLLEENKIKLTHGQRFPKAYHHALHSLFLAITIVESQYAEIPKE
                     SRNVNYSLASFLKCCLTLMDRGFVFNLINDYISGFSPKDPKVLAEYKFEFLQTICNHE
                     HYIPLNLPMAFAKPKLQRVQDSNLEYSLSDEYCKHHFLVGLLLRETSIALQDNYEIRY
                     TAISVIKNLLIKHAFDTRYQHKNQQAKIAQLYLPFVGLLLENIQRLAGRDTLYSCAAM
                     PNSASRDEFPCGFTSPANRGSLSTDKDTAYGSFQNGHGIKREDSRGSLIPEGATGFPD
                     QGNTGENTRQSSTRSSVSQYNRLDQYEIRSLLMCYLYIVKMISEDTLLTYWNKVSPQE
                     LINILILLEVCLFHFRYMGKRNIARVHDAWLSKHFGIDRKSQTMPALRNRSGVMQARL
                     QHLSSLESSFTLNHSSTTTEADIFHQALLEGNTATEVSLTVLDTISFFTQCFKTHFLN
                     NDGHNPLMKKVFDIHLAFLKNGQSEVSLKHVFASLRAFISKFPSAFFKGRVNMCAAFC
                     YEVLKCCTSKISSTRNEASALLYLLMRNNFEYTKRKTFLRTHLQIIIAVSQLIADVAL
                     SGGSRFQESLFIINNFANSDRPMLARAFPAEVKDLTKRIRTVLMATAQMKEHEKDPEM
                     LIDLQYSLAKSYASTPELRKTWLDSMAKIHVKNGDFSEAAMCYVHVAALVAEFLHRKK
                     LFPNGCSAFKKITPNIDEEGAMKEDAGMMDVHYSEEVLLELLEQCVNGLWKAERYEII
                     SEISKLIGPIYENRREFENLTQVYRTLHGAYTKILEVMHTKKRLLGTFFRVAFYGQSF
                     FEEEDGKEYIYKEPKLTGLSEISLRLVKLYGEKFGTENVKIIQDSDKVNAKELDPKYA
                     HIQVTYVKPYFDDKELTERKTEFERNHNISRFVFEAPYTLSGKKQGCIEEQCKRRTIL
                     TTSNSFPYVKKRIPINCEQQINLKPIDGATDEIKDKTAELQKLCSSTDVDMIQLQLKL
                     QGWVSVQVNAGPLAYARAFLNDSQASKYPPKKVSELKDMFRKFIQACSIALELNERLI
                     KEDQVEYHEGLKSNFRDMVKELSDIIHEQILQEDTMHSPWMSNTLHVFCAISGTSSDR
                     GYGSPRYAEV"
BASE COUNT     1359 a    851 c    908 g   1273 t
ORIGIN

  Query Match           53.8%;  Score 269.6;  DB 6;  Length 4391;
  Best Local Similarity  71.2%;  Pred. No. 4.8e-61;
  Matches  356;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db    2379 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 2438

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db    2439 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 2498

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           || || | ||| || |||||| | | |||   |   |    |||  || | || || ||
Db    2499 AATAATTTTGCAAATAGTGACAGACCTATGTTGGCAAGAGCCTTTCCCGCAGAAGTCAAA 2558

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||||| ||||| || || ||||| || || |||||||| ||||| ||||||||||||||||
Db    2559 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 2618

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db    2619 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 2678

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db    2679 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 2738

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
```

```
          | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db   2739 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 2798

Qy    422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
          || |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db   2799 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 2858

Qy    482 GAGGAGGCCTCCATGATGGA 501
          || || |   | ||||  ||
Db   2859 GAAGAAGGAGCAATGAAAGA 2878


RESULT 14
AX173118
LOCUS       AX173118                4393 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 97 from Patent WO0142294.
ACCESSION   AX173118
VERSION     AX173118.1  GI:14598012
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4393)
  AUTHORS   Lu,P., Garman,J.D. and Candia,A.F.
  TITLE     Clasp-4 transmembrane protein
  JOURNAL   Patent: WO 0142294-A 97 14-JUN-2001;
            Arbor Vita Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .4393
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="polynucleotide fragment"
BASE COUNT     1359 a    852 c    910 g   1272 t
ORIGIN

  Query Match           53.8%;  Score 269.6;  DB 6;  Length 4393;
  Best Local Similarity 71.2%;  Pred. No. 4.8e-61;
  Matches  356;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy      2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
          || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db   2379 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 2438

Qy     62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
          |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db   2439 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 2498

Qy    122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
          || || | ||| || |||||| | | |||   |   |    |||  || | || || ||
Db   2499 AATAATTTTGCAAATAGTGACAGACCTATGTTGGCAAGAGCCTTTCCCGCAGAAGTCAAA 2558

Qy    182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
          ||||| ||||| || || ||||| || || |||||||| |||| ||||||||||||||||
Db   2559 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 2618

Qy    242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
          || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db   2619 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 2678

Qy    302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
          || |||||||||||| || ||||| || |||||||||| ||| ||||| ||||||| |||
Db   2679 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 2738

Qy    362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db   2739 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 2798

Qy    422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
          || |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db   2799 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 2858

Qy    482 GAGGAGGCCTCCATGATGGA 501
          || || |   | ||||  ||
Db   2859 GAAGAAGGAGCAATGAAAGA 2878


RESULT 15
AK054649
LOCUS       AK054649                2299 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ30087 fis, clone BNGH41000003, moderately
            similar to R.norvegicus trg mRNA.
ACCESSION   AK054649
VERSION     AK054649.1  GI:16549234
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens neuroglioma cell_line:H4 cDNA to mRNA,
            clone_lib:BNGH41 clone:BNGH41000003.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K.,
            Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
            Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2299)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES             Location/Qualifiers
     source          1. .2299
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BNGH41000003"
                     /cell_line="H4"
                     /cell_type="neuroglioma"
                     /clone_lib="BNGH41"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      780 a    392 c    471 g    656 t
ORIGIN

  Query Match           52.6%;  Score 263.4;  DB 9;  Length 2299;
  Best Local Similarity 71.7%;  Pred. No. 2.1e-59;
  Matches  359;  Conservative    0;  Mismatches  141;  Indels    1;  Gaps    1;

Qy      2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
          || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db    287 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 346

Qy     62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
          |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db    347 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 406

Qy    122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
          || || | ||| || |||||| | | ||| |||   ||   |||  || | || || ||
Db    407 AATAATTTTGCAAATAGTGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAA 466

Qy    182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
          ||||| ||||| || || ||||| || || |||||||| ||||| ||||||||||||||||
Db    467 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 526
```

```
Qy    242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
          || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db    527 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 586

Qy    302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
          || |||||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db    587 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 646

Qy    362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATC-TCAC 420
           | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | || |||
Db    647 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 706

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
              |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db    707 CGAAAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGA 766

Qy    481 CGAGGAGGCCTCCATGATGGA 501
           || || |   | ||||  ||
Db    767 TGAAGAAGGAGCAATGAAAGA 787
```

Search completed: February 7, 2003, 08:03:39
Job time : 1021.66 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:36:04 ; Search time 128.085 Seconds
(without alignments)
8808.598 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaaagaagtcct.....gaggaggcctccatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	501	100.0	3614	24	ABK84967	DNA encoding cadhe
2	501	100.0	3642	24	ABK84971	DNA encoding cadhe
3	501	100.0	3705	24	ABK84968	DNA encoding cadhe
4	501	100.0	4806	24	ABK84964	DNA encoding cadhe
5	501	100.0	4807	21	AAC87972	Human CLASP-2 nucl
6	501	100.0	4807	21	AAC87973	Human CLASP-2A nuc
7	501	100.0	4807	24	ABK84966	DNA encoding cadhe
8	501	100.0	4807	24	ABK84973	DNA encoding cadhe
9	501	100.0	4898	21	AAC87974	Preliminary CLASP-
10	501	100.0	4898	21	AAC87975	Preliminary CLASP-
11	501	100.0	4898	21	AAC87976	Preliminary CLASP-
12	501	100.0	4898	21	AAC87977	Preliminary CLASP-
13	501	100.0	4898	21	AAC87978	Preliminary CLASP-
14	501	100.0	4898	21	AAC87979	Preliminary CLASP-
15	501	100.0	4898	21	AAC87980	Preliminary CLASP-
16	501	100.0	4898	21	AAC87981	Preliminary CLASP-
17	501	100.0	4898	24	ABK84992	DNA encoding cadhe
18	501	100.0	4898	24	ABK84993	DNA encoding cadhe
19	501	100.0	4898	24	ABK84994	DNA encoding cadhe
20	501	100.0	4898	24	ABK84995	DNA encoding cadhe
21	501	100.0	4898	24	ABK84996	DNA encoding cadhe
22	501	100.0	4898	24	ABK84997	DNA encoding cadhe
23	501	100.0	4898	24	ABK84998	DNA encoding cadhe
24	501	100.0	4898	24	ABK84999	DNA encoding cadhe
25	501	100.0	6791	24	ABK85003	DNA encoding cadhe
26	499.4	99.7	3678	23	ABV21145	Human prostate exp
27	499.4	99.7	3899	22	AAF32675	Human cDNA encodin
28	426.4	85.1	5862	24	ABK84970	DNA encoding cadhe
29	426.4	85.1	6816	21	AAC74524	Human ORFX ORF79 p
30	426.4	85.1	7506	22	AAD19118	Angiogenesis assoc
31	422	84.2	5048	24	ABK84965	DNA encoding cadhe
32	274.4	54.8	2610	22	AAF32647	Human cDNA encodin
33	274.4	54.8	6454	22	AAS08334	Human cDNA encodin
34	269.6	53.8	3472	22	AAS08358	Human cDNA encodin
35	269.6	53.8	4393	22	AAS08357	Human cDNA encodin
36	221.2	44.2	5214	21	AAC87968	Mouse CLASP-1 nucl
37	221.2	44.2	5214	21	AAA14824	DNA encoding a mur
38	216.4	43.2	5688	21	AAC87969	Human CLASP relate
39	216.4	43.2	5688	21	AAA14825	DNA encoding a hum
40	216.4	43.2	7277	24	AAS18951	Human cDNA encodin
41	171.4	34.2	5955	23	ABL11149	Drosophila melanog
C 42	171.4	34.2	8942	23	ABL11148	Drosophila melanog
43	168.2	33.6	1476	24	ABA92670	Human NOV9 encodin
44	136	27.1	2120	22	AAK72038	Human immune/haema
45	130.6	26.1	15057	22	AAK72037	Human immune/haema

ALIGNMENTS

RESULT 1
ABK84967
ID ABK84967 standard; cDNA; 3614 BP.
XX
AC ABK84967;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #2.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.

XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61673.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 3614 BP; 1009 A; 834 C; 874 G; 897 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 3614;
Best Local Similarity 100.0%; Pred. No. 9.9e-142;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1307 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 1366

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1367 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 1426

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1427 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 1486

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1487 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 1546

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1547 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 1606

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1607 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 1666

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1667 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 1726

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1727 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 1786

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    1787 CGAGGAGGCCTCCATGATGGA 1807
```

RESULT 2
ABK84971
ID ABK84971 standard; cDNA; 3642 BP.
XX
AC ABK84971;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #6.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61677.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that

```
CC   binds CLASP-2 polypeptide.  An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g., agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. CLASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC   PCR primers of the invention.
XX
SQ   Sequence 3642 BP; 1005 A; 899 C; 877 G; 861 T; 0 other;

  Query Match           100.0%;  Score 501;  DB 24;  Length 3642;
  Best Local Similarity  100.0%;  Pred. No. 9.9e-142;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2028 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2087

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2088 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2147

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2148 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2207

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2208 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2267

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2268 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2327

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2328 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2387

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2388 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2447

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2448 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2507

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2508 CGAGGAGGCCTCCATGATGGA 2528


RESULT 3
ABK84968
ID   ABK84968 standard; cDNA; 3705 BP.
XX
AC   ABK84968;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) isoform #3.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200231117-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
PA   (ARBO-) ARBOR VITA CORP.
PA   (GARM/) GARMAN J D.
PA   (CAND/) CANDIA A F.
XX
PI   Lu PS;
XX
DR   WPI; 2002-416861/44.
DR   P-PSDB; ABG61674.
XX
PT   New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT   an immune response, and for treating multiple sclerosis, rheumatoid
PT   arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT   and sepsis -
XX
PS   Disclosure; Figure 3A; 245pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising an amino
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide.  An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
```

CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 3705 BP; 1017 A; 911 C; 899 G; 878 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 3705;
Best Local Similarity 100.0%; Pred. No. 1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2022 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2081

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pb    2082 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2141

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2142 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2201

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2202 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2261

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2262 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2321

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2322 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2381

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2382 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2441

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2442 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2501

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2502 CGAGGAGGCCTCCATGATGGA 2522
```

RESULT 4
ABK84964
ID ABK84964 standard; cDNA; 4806 BP.
XX
AC ABK84964;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP).
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61670.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 1; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018

```
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC   PCR primers of the invention.
XX
SQ   Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 24;  Length 4806;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2499 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2558

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2559 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2618

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2619 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2678

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2679 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2738

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2739 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2798

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2799 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2858

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2859 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2918

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2919 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2978

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2979 CGAGGAGGCCTCCATGATGGA 2999


RESULT 5
AAC87972
ID   AAC87972 standard; cDNA; 4807 BP.
XX
AC   AAC87972;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Human CLASP-2 nucleotide sequence.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
DR   P-PSDB; AAB36527.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Example 1; Fig 1; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence encodes human CLASP-2, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4807;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2559

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2619

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2679

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2739

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2799

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2800 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2859

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2919

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2979

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2980 CGAGGAGGCCTCCATGATGGA 3000


RESULT 6
AAC87973
ID   AAC87973 standard; cDNA; 4807 BP.
XX
AC   AAC87973;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Human CLASP-2A nucleotide sequence.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
DR   P-PSDB; AAB36528.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Example 1; Fig 2B; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence encodes human CLASP-2A, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4807;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2559

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2619

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2679

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2739

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2799

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2800 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2859

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2919

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2979

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2980 CGAGGAGGCCTCCATGATGGA 3000


RESULT 7
ABK84966
ID   ABK84966 standard; cDNA; 4807 BP.
XX
AC   ABK84966;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) isoform #1.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
```

KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61672.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2559

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2619

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2679

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2739

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2799

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2800 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2859

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2919

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2979

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2980 CGAGGAGGCCTCCATGATGGA 3000
```

RESULT 8
ABK84973
ID ABK84973 standard; DNA; 4807 BP.
XX
AC ABK84973;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.

PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61686.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Example 4; Figure 6A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences
CC and PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2559

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2619

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2679

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2739

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2799

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2800 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2859

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 2919

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 2979

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    2980 CGAGGAGGCCTCCATGATGGA 3000
```

RESULT 9
AAC87974
ID AAC87974 standard; cDNA; 4898 BP.
XX
AC AAC87974;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #1.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PSDB; AAB36529.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -

```
XX
PS   Disclosure; Fig 10A; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    3071 CGAGGAGGCCTCCATGATGGA 3091


RESULT 10
AAC87975
ID   AAC87975 standard; cDNA; 4898 BP.
XX
AC   AAC87975;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #2.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10B; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy     61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy    121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy    181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy    241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy    301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy    361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   3071 CGAGGAGGCCTCCATGATGGA 3091
```

RESULT 11
AAC87976
ID AAC87976 standard; cDNA; 4898 BP.
XX
AC AAC87976;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #3.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10C; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy   61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy   121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy   181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy   241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy   301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy   361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy   421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

```
Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   3071 CGAGGAGGCCTCCATGATGGA 3091


RESULT 12
AAC87977
ID   AAC87977 standard; cDNA; 4898 BP.
XX
AC   AAC87977;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #4.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10D; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy     61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy    121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy    181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy    241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy    301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy    361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   3071 CGAGGAGGCCTCCATGATGGA 3091


RESULT 13
AAC87978
ID   AAC87978 standard; cDNA; 4898 BP.
XX
AC   AAC87978;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #5.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
```

```
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10E; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-141;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy     481 CGAGGAGGCCTCCATGATGGA 501
           |||||||||||||||||||||
Db    3071 CGAGGAGGCCTCCATGATGGA 3091


RESULT 14
AAC87979
ID   AAC87979 standard; cDNA; 4898 BP.
XX
AC   AAC87979;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #6.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10F; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
```

CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy     481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db    3071 CGAGGAGGCCTCCATGATGGA 3091
```

RESULT 15
AAC87980
ID AAC87980 standard; cDNA; 4898 BP.
XX
AC AAC87980;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #7.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10G; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1379 A; 1134 C; 1166 G; 1219 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2591 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 2650

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2651 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 2710

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2711 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 2770

Qy    181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2771 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 2830

Qy    241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2831 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 2890

Qy    301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2891 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 2950

Qy    361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2951 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 3010

Qy    421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3011 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 3070

Qy    481 CGAGGAGGCCTCCATGATGGA 501
          |||||||||||||||||||||
Db   3071 CGAGGAGGCCTCCATGATGGA 3091
```

Search completed: February 7, 2003, 07:08:07
Job time : 158.085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 ; Search time 28.6858 Seconds
(without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaaagaagtcct.....gaggaggcctccatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		43	8.6	4722	4	US-08-979-608A-14	Sequence 14, Appl
2		37.4	7.5	1432	1	US-07-914-282D-3	Sequence 3, Appli
3		37.4	7.5	1432	1	US-08-276-887A-3	Sequence 3, Appli
4		37.4	7.5	1432	5	PCT-US93-02460-3	Sequence 3, Appli
5		36.4	7.3	1111	2	US-08-997-080-169	Sequence 169, App
6		36.4	7.3	1111	2	US-08-997-362-169	Sequence 169, App
7		36.4	7.3	1111	4	US-09-095-855-169	Sequence 169, App
8		36.4	7.3	1111	4	US-09-324-542-169	Sequence 169, App
9		36.4	7.3	1111	4	US-09-205-426-169	Sequence 169, App
10		36.4	7.3	1341	2	US-08-997-080-93	Sequence 93, Appl
11		36.4	7.3	1341	2	US-08-997-362-93	Sequence 93, Appl
12		36.4	7.3	1341	3	US-08-873-970-93	Sequence 93, Appl
13		36.4	7.3	1341	4	US-09-095-855-93	Sequence 93, Appl
14		36.4	7.3	1341	4	US-09-324-542-93	Sequence 93, Appl
15		36.4	7.3	1341	4	US-09-205-426-93	Sequence 93, Appl
16	c	35	7.0	1854	5	PCT-US94-01101-1	Sequence 1, Appli
17		35	7.0	4697	4	US-08-979-608A-17	Sequence 17, Appl
18	c	33.6	6.7	28171	4	US-08-961-527-22	Sequence 22, Appl
19		32.6	6.5	4972	3	US-09-035-648-17	Sequence 17, Appl
20		32.6	6.5	4972	4	US-09-001-951-17	Sequence 17, Appl
21		32.6	6.5	4972	4	US-08-818-829-17	Sequence 17, Appl
22		32.2	6.4	1136	1	US-08-106-981-5	Sequence 5, Appli
23		32.2	6.4	4403765	4	US-09-103-840A-2	Sequence 2, Appli
24		32.2	6.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
25	c	32	6.4	2419	4	US-09-245-041-8	Sequence 8, Appli
26	c	32	6.4	8827	4	US-09-245-041-1	Sequence 1, Appli
27	c	32	6.4	62804	4	US-09-800-960-3	Sequence 3, Appli
28		31.8	6.3	4443	4	US-09-425-453A-1	Sequence 1, Appli
29		31.8	6.3	4443	4	US-09-425-453A-3	Sequence 3, Appli
30		31.8	6.3	4443	4	US-09-425-453A-5	Sequence 5, Appli
31		31.8	6.3	4443	4	US-09-425-453A-7	Sequence 7, Appli
32		31.8	6.3	4443	4	US-09-425-453A-9	Sequence 9, Appli
33		31.8	6.3	4443	4	US-09-425-453A-11	Sequence 11, Appl
34		31.8	6.3	4443	4	US-09-425-453A-13	Sequence 13, Appl
35		31.8	6.3	4443	4	US-09-425-453A-15	Sequence 15, Appl
36		31.8	6.3	4443	4	US-09-425-453A-17	Sequence 17, Appl
37		31.8	6.3	4443	4	US-09-425-453A-19	Sequence 19, Appl
38		31.8	6.3	4560	4	US-09-256-703-1	Sequence 1, Appli
39		31.8	6.3	5635	1	US-08-136-742A-3	Sequence 3, Appli
40		31.8	6.3	5635	3	US-09-248-026-3	Sequence 3, Appli
41		31.8	6.3	5635	5	PCT-US93-11667-3	Sequence 3, Appli
42		31.8	6.3	6126	2	US-08-951-912-3	Sequence 3, Appli
43		31.8	6.3	6126	4	US-09-174-077-3	Sequence 3, Appli
44		31.8	6.3	6129	1	US-07-637-621-1	Sequence 1, Appli
45		31.8	6.3	6129	1	US-08-136-742A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-979-608A-14
; Sequence 14, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...1731
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-08-979-608A-14

Query Match 8.6%; Score 43; DB 4; Length 4722;
Best Local Similarity 50.7%; Pred. No. 0.0011;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy      84 TTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACC 143
           | | |||  |||   ||||||||| ||| |   ||  | ||  |    |   |  |||
Db    1067 TGGTGGACGCCAAGCTCCAGCAGGCCCAGGAGATGCTGAAGGAGGCAGAGGAGCGGCACC 1126

Qy     144 GGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCA 203
            ||     ||| ||     | |       | ||| ||  |   |  |   |    |  |
Db    1127 AGCGGGAGAAGGACTTTCTCCTGAAGGAGGCCGTGGAGTCCCAGAGGATGTGCGAGCTGA 1186

Qy     204 CGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGG 263
           |  || |  ||  |||  || || | || ||| ||      || ||||||  | | | ||
Db    1187 TGAAGCAACAGGAGACCCACCTGAAGCAGCAGCTTGCCCTATACACAGAGAAGTTTGAGG 1246

Qy     264 ACCTCCAGTACAGCCTGGCCAAA 286
           |  ||||| |||  ||  |||||
Db    1247 AGTTCCAGAACACTCTTTCCAAA 1269
```

RESULT 2
US-07-914-282D-3
; Sequence 3, Application US/07914282D
; Patent No. 5364787
; GENERAL INFORMATION:
; APPLICANT: Orser, Cindy S. and Xun, Luying
; TITLE OF INVENTION: GENES AND ENZYMES INVOLVED IN
; TITLE OF INVENTION: THE MICROBIAL DEGRADATION OF
; TITLE OF INVENTION: PENTACHLOROPHENOL
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ramon A. Klitzke II
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1/PC Gene
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,282D
; FILING DATE: July 13, 1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/856,015
; FILING DATE: March 23, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramon A. Klitzke II
; REGISTRATION NUMBER: 30,188
; REFERENCE/DOCKET NUMBER: 2815-36746
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1432 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double-stranded
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Flavobacterium sp. Strain ATCC 39723
US-07-914-282D-3

Query Match 7.5%; Score 37.4; DB 1; Length 1432;
Best Local Similarity 48.0%; Pred. No. 0.036;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           || | ||   |||      |||||  |  |  | ||   ||  | | ||| ||| | | |
Db     819 CATCTACCAGGCCGCGCACGACCGCATCGTGGAACATGGCAATTGCGCCTATGACGCGGA 878

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
               |  || |   |  | |    ||| || |  |  || | |  || | | ||||
Db     879 CACGGTGGCCCAGGCGGAAGTCGACCTGCAAAAGCGCCTCGACGAACTGGACGTGCATCT 938

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           |   |||     | |  | | || |  | | |||||| | ||| |   |      | |
Db     939 GGCGGACAAGCCGTTCATCGCGGGCAGCAACTACAGCATCGCCGACATCATGTGGACCGT 998

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATC 343
             |    |    |||    |  |||| |||  | |||| |||||
Db     999 CCTGCTGGCGCGGATCGAGATGCTCAACATGACGGCCTGGATC 1041
```

RESULT 3
US-08-276-887A-3
; Sequence 3, Application US/08276887A
; Patent No. 5512478
; GENERAL INFORMATION:
; APPLICANT: Orser, Cindy S. and Xun, Luying
; TITLE OF INVENTION: GENES AND ENZYMES INVOLVED
; TITLE OF INVENTION: IN THE MICROBIAL
; TITLE OF INVENTION: DEGRADATION OF
; TITLE OF INVENTION: PENTACHLOROPHENOL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ramon A. Klitzke II
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1/PC Gene
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,887A
; FILING DATE: July 18, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/914,282
; FILING DATE: July 13, 1992
; APPLICATION NUMBER: 07/856,015
; FILING DATE: March 23, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramon A. Klitzke II
; REGISTRATION NUMBER: 30,188
; REFERENCE/DOCKET NUMBER: 2815-36746
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1432 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double-stranded
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; DESCRIPTION:

```
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Flavobacterium sp. Strain
;      ORGANISM:  ATCC 39723
US-08-276-887A-3

  Query Match             7.5%;  Score 37.4;  DB 1;  Length 1432;
  Best Local Similarity   48.0%;  Pred. No. 0.036;
  Matches  107;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           || | ||   |||      ||||| |  | | ||   ||  | | ||| ||| | | |
Db     819 CATCTACCAGGCCGCGCACGACCGCATCGTGGAACATGGCAATTGCGCCTATGACGCGGA 878

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
               | || |   | | |    ||| || | | || | |  || | | |||| 
Db     879 CACGGTGGCCCAGGCGGAAGTCGACCTGCAAAAGCGCCTCGACGAACTGGACGTGCATCT 938

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           |  |||    | |  | | || |  | | |||||| | ||| |   |      | |
Db     939 GGCGGACAAGCCGTTCATCGCGGGCAGCAACTACAGCATCGCCGACATCATGTGGACCGT 998

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATC 343
             |    |   |||    |  |||| |||  | |||| |||||
Db     999 CCTGCTGGCGCGGATCGAGATGCTCAACATGACGGCCTGGATC 1041


RESULT 4
PCT-US93-02460-3
; Sequence 3, Application PC/TUS9302460
;  GENERAL INFORMATION:
;    APPLICANT:  Orser, Cindy S. and Xun, Luying
;    TITLE OF INVENTION:  GENES AND ENZYMES INVOLVED IN THE
;    TITLE OF INVENTION:  MICROBIAL DEGRADATION OF PENTACHLOROPHENOL
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William D. Noonan, M.D.
;      STREET:  One World Trade Center
;      STREET:  121 S.W. Salmon Street
;      STREET:  Suite 1600
;      CITY:  Portland
;      STATE:  Oregon
;      COUNTRY:  United States of America
;      ZIP:  97204
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5-inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect 5.1/PC Gene
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02460
;      FILING DATE:  19930319
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/856,015
;      FILING DATE:  March 23, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/914,282
;      FILING DATE:  July 13, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  William D. Noonan, M.D.
;      REGISTRATION NUMBER:  30878
;      REFERENCE/DOCKET NUMBER:  2815-36746-WDN
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (503) 226-7391
;      TELEFAX:  (503) 228-9446
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1432 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Double-stranded
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Genomic DNA
;      DESCRIPTION:
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Flavobacterium sp. Strain ATCC 39723
PCT-US93-02460-3

  Query Match             7.5%;  Score 37.4;  DB 5;  Length 1432;
  Best Local Similarity   48.0%;  Pred. No. 0.036;
  Matches  107;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           || | ||   |||      |||||  |  | | ||   ||  | | ||| ||| | | |
Db     819 CATCTACCAGGCCGCGCACGACCGCATCGTGGAACATGGCAATTGCGCCTATGACGCGGA 878

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
               | || |   | | |    ||| || | | || | |  || | | |||| 
Db     879 CACGGTGGCCCAGGCGGAAGTCGACCTGCAAAAGCGCCTCGACGAACTGGACGTGCATCT 938

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           |  |||    | |  | | || |  | | |||||| | ||| |   |      | |
Db     939 GGCGGACAAGCCGTTCATCGCGGGCAGCAACTACAGCATCGCCGACATCATGTGGACCGT 998

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATC 343
             |    |   |||    |  |||| |||  | |||| |||||
Db     999 CCTGCTGGCGCGGATCGAGATGCTCAACATGACGGCCTGGATC 1041


RESULT 5
US-08-997-080-169
; Sequence 169, Application US/08997080
; Patent No. 5968524
;  GENERAL INFORMATION:
;    APPLICANT:  WATSON, JAMES D.
;    APPLICANT:  TAN, PAUL L.J.
;    TITLE OF INVENTION:  METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,080
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  169:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1111 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

US-08-997-080-169

Query Match 7.3%; Score 36.4; DB 2; Length 1111;
Best Local Similarity 46.2%; Pred. No. 0.066;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```
Qy    100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
          | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db    805 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 864

Qy    160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
          |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db    865 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 924

Qy    220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
           ||  |||    |  |    |        | ||| ||| || | |||| |||   |
Db    925 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 984

Qy    280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
          |||     || | |     |||| ||| ||| |   |||    || |  ||  | | |
Db    985 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1044

Qy    340 GATCCATGTCAAAAATGGCGAT 361
            |  |     | || ||||
Db   1045 CTGACGCGGTGGTAGTGCCGAT 1066
```

RESULT 6
US-08-997-362-169
; Sequence 169, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 169:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1111 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-362-169

Query Match 7.3%; Score 36.4; DB 2; Length 1111;
Best Local Similarity 46.2%; Pred. No. 0.066;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

Qy 100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
Db 805 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 864

Qy 160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
Db 865 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 924

Qy 220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
Db 925 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 984

Qy 280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
Db 985 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1044

Qy 340 GATCCATGTCAAAAATGGCGAT 361
Db 1045 CTGACGCGGTGGTAGTGCCGAT 1066

RESULT 7
US-09-095-855-169
; Sequence 169, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:

```
;  INFORMATION FOR SEQ ID NO:  169:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1111 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-095-855-169

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1111;
  Best Local Similarity   46.2%;  Pred. No. 0.066;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     805 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 864

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db     865 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 924

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db     925 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 984

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| |||  |   |||    || |  ||  | | |
Db     985 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1044

Qy     340 GATCCATGTCAAAAATGGCGAT 361
           |  |     | ||  ||||
Db    1045 CTGACGCGGTGGTAGTGCCGAT 1066


RESULT 8
US-09-324-542-169
; Sequence 169, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L.J.
;  APPLICANT: Prestidge, Ross
;  TITLE OF INVENTION: Methods and Compounds for the Treatment
;  TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
;  FILE REFERENCE: 11000.1007c1
;  CURRENT APPLICATION NUMBER: US/09/324,542
;  CURRENT FILING DATE: 1999-06-02
;  EARLIER APPLICATION NUMBER: US 08/997,080
;  EARLIER FILING DATE: 1997-12-23
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 169
;   LENGTH: 1111
;   TYPE: DNA
;   ORGANISM: Mycobacterium vaccae
US-09-324-542-169

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1111;
  Best Local Similarity   46.2%;  Pred. No. 0.066;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     805 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 864

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db     865 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 924

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db     925 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 984

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| |||  |   |||    || |  ||  | | |
Db     985 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1044

Qy     340 GATCCATGTCAAAAATGGCGAT 361
           |  |     | ||  ||||
Db    1045 CTGACGCGGTGGTAGTGCCGAT 1066


RESULT 9
US-09-205-426-169
; Sequence 169, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L. J.
;  TITLE OF INVENTION: Compounds and Methods for Treatment and
;  TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
;  FILE REFERENCE: 11000.1002c4
;  CURRENT APPLICATION NUMBER: US/09/205,426
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: 09/095,855
;  EARLIER FILING DATE: 1998-06-11
;  EARLIER APPLICATION NUMBER: 08/997,362
;  EARLIER FILING DATE: 1997-12-23
;  EARLIER APPLICATION NUMBER: 08/873,970
;  EARLIER FILING DATE: 1997-06-12
;  EARLIER APPLICATION NUMBER: 08/705,347
;  EARLIER FILING DATE: 1996-08-29
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 169
;   LENGTH: 1111
;   TYPE: DNA
;   ORGANISM: Mycobacterium vaccae
US-09-205-426-169

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1111;
  Best Local Similarity   46.2%;  Pred. No. 0.066;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     805 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 864

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db     865 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 924

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db     925 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 984

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| |||  |   |||    || |  ||  | | |
Db     985 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1044

Qy     340 GATCCATGTCAAAAATGGCGAT 361
           |  |     | ||  ||||
Db    1045 CTGACGCGGTGGTAGTGCCGAT 1066


RESULT 10
US-08-997-080-93
; Sequence 93, Application US/08997080
; Patent No. 5968524
;  GENERAL INFORMATION:
;    APPLICANT:  WATSON, JAMES D.
;    APPLICANT:  TAN, PAUL L.J.
;    TITLE OF INVENTION:  METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
```

```
;       CITY:  Seattle
;       STATE:  WA
;       COUNTRY:  USA
;       ZIP:  98121
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/997,080
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sleath, Janet
;       REGISTRATION NUMBER:  37,007
;       REFERENCE/DOCKET NUMBER:  11000.1007
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  206-269-0565
;       TELEFAX:  206-269-0563
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  93:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1341 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Genomic DNA
US-08-997-080-93

  Query Match             7.3%;  Score 36.4;  DB 2;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |      |   || |||        ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     |||| ||| ||| |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
               |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


RESULT 11
US-08-997-362-93
; Sequence 93, Application US/08997362
; Patent No. 5985287
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,362
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/873,970
;      FILING DATE:  June 12, 1997
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/705,347
;      FILING DATE:  August 29, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-997-362-93

  Query Match             7.3%;  Score 36.4;  DB 2;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |      |   || |||        ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     |||| ||| ||| |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
               |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


RESULT 12
US-08-873-970-93
; Sequence 93, Application US/08873970
; Patent No. 6001361
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
```

```
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/873,970
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-873-970-93

  Query Match             7.3%;  Score 36.4;  DB 3;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| ||| |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
           |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


RESULT 13
US-09-095-855-93
; Sequence 93, Application US/09095855
; Patent No. 6160093
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  Compounds and Methods for
;    TITLE OF INVENTION:  Treatment and Diagnosis of Mycobacterial Infections
;    NUMBER OF SEQUENCES:  208
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/095,855
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;      APPLICATION NUMBER:  08/873,970
;      FILING DATE:  12-JUN-1997
;      APPLICATION NUMBER:  08/997,362
;      FILING DATE:  23-DEC-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-095-855-93

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |        | ||| ||| || | ||||  |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| ||| |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
           |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


RESULT 14
US-09-324-542-93
; Sequence 93, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
```

```
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L.J.
;  APPLICANT: Prestidge, Ross
;  TITLE OF INVENTION: Methods and Compounds for the Treatment
;  TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
;  FILE REFERENCE: 11000.1007c1
;  CURRENT APPLICATION NUMBER: US/09/324,542
;  CURRENT FILING DATE: 1999-06-02
;  EARLIER APPLICATION NUMBER: US 08/997,080
;  EARLIER FILING DATE: 1997-12-23
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
;   LENGTH: 1341
;   TYPE: DNA
;   ORGANISM: Mycobacterium vaccae
US-09-324-542-93

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |         | ||| ||| || | |||| |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| |||  |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
             |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


RESULT 15
US-09-205-426-93
; Sequence 93, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L. J.
;  TITLE OF INVENTION: Compounds and Methods for Treatment and
;  TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
;  FILE REFERENCE: 11000.1002c4
;  CURRENT APPLICATION NUMBER: US/09/205,426
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: 09/095,855
;  EARLIER FILING DATE: 1998-06-11
;  EARLIER APPLICATION NUMBER: 08/997,362
;  EARLIER FILING DATE: 1997-12-23
;  EARLIER APPLICATION NUMBER: 08/873,970
;  EARLIER FILING DATE: 1997-06-12
;  EARLIER APPLICATION NUMBER: 08/705,347
;  EARLIER FILING DATE: 1996-08-29
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
;   LENGTH: 1341
;   TYPE: DNA
;   ORGANISM: Mycobacterium vaccae
US-09-205-426-93

  Query Match             7.3%;  Score 36.4;  DB 4;  Length 1341;
  Best Local Similarity   46.2%;  Pred. No. 0.072;
  Matches  121;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 CCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACAC 159
           | ||  ||   |   |  |||| || || | ||||||   | |  |||  |   |  |||
Db     979 CGAGGCGTGGATCGACTACATCTACGACCGAGCCAACTACGCCAAGCTGGTCGCGTTCAC 1038

Qy     160 CAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCAC 219
           |   |||    | |   |   |||| | ||| |     |   || |||       ||  |
Db    1039 CCAGTTCGTGCCCGCACTCTCGGACATGACCGACGAACTCGCCAAGGTCGATCCTGCATC 1098

Qy     220 CGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCT 279
            ||  |||    |  |    |         | ||| ||| || | |||| |||   |
Db    1099 GGCGGAGAACCCGCTGATCAACCCGTCGGCCGAGGTGCAGGCGAACCTGAAGTCGTGGGC 1158

Qy     280 GGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAG 339
           |||     || | |     ||||  ||| |||  |   |||    || |  ||  | | |
Db    1159 GGCACTGACCGACGAGCAGACGCAGGAGTTCAACACTGCGTACGCCGCCGTCACCGGCGG 1218

Qy     340 GATCCATGTCAAAAATGGCGAT 361
             |  |     | || ||||
Db    1219 CTGACGCGGTGGTAGTGCCGAT 1240


Search completed: February  7, 2003, 09:05:47
Job time : 43.6858 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.3549 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaaagaagtcct.....gaggaggcctccatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	274.4	54.8	6454	10	US-09-736-969A-1	Sequence 1, Appli
2	269.6	53.8	4391	10	US-09-736-969A-7	Sequence 7, Appli
3	130.6	26.1	483	10	US-09-864-761-13960	Sequence 13960, A
4	116.4	23.2	2149	9	US-09-736-968A-59	Sequence 59, Appl
5	116.4	23.2	6372	9	US-09-736-968A-1	Sequence 1, Appli
6	115.4	23.0	212	10	US-09-864-761-30524	Sequence 30524, A
7	83.2	16.6	4026	10	US-09-736-960-3	Sequence 3, Appli
8	83.2	16.6	7215	10	US-09-736-960-1	Sequence 1, Appli
9	79.4	15.8	427	9	US-09-796-692-5465	Sequence 5465, Ap
10	69	13.8	444	9	US-09-796-692-3710	Sequence 3710, Ap
11	57.4	11.5	271	9	US-09-736-968A-98	Sequence 98, Appl
12	57.4	11.5	512	10	US-09-864-761-7634	Sequence 7634, Ap
13	50	10.0	194	10	US-09-864-761-24339	Sequence 24339, A
14	48.8	9.7	266	10	US-09-736-960-76	Sequence 76, Appl
15	48.8	9.7	66686	10	US-09-736-960-86	Sequence 86, Appl
16	46.2	9.2	379	9	US-10-046-935-2206	Sequence 2206, Ap
17	46.2	9.2	379	9	US-09-878-178-2206	Sequence 2206, Ap
18	45.4	9.1	181	9	US-09-736-968A-97	Sequence 97, Appl
19	43	8.6	4722	9	US-09-976-740-14	Sequence 14, Appl
20	43	8.6	4722	10	US-09-962-055-14	Sequence 14, Appl
21	43	8.6	4722	12	US-10-023-529-14	Sequence 14, Appl
22	43	8.6	4722	12	US-10-023-523-14	Sequence 14, Appl
23	36.4	7.3	1111	9	US-10-051-643-169	Sequence 169, App
24	36.4	7.3	1111	9	US-09-880-505-169	Sequence 169, App
25	36.4	7.3	1341	9	US-10-051-643-93	Sequence 93, Appl
26	36.4	7.3	1341	9	US-09-880-505-93	Sequence 93, Appl
c 27	35.4	7.1	292	10	US-09-815-242-872	Sequence 872, App
28	35.4	7.1	405	10	US-09-815-242-3815	Sequence 3815, Ap
29	35.4	7.1	435	10	US-09-815-242-6463	Sequence 6463, Ap
c 30	35.4	7.1	558	10	US-09-815-242-112	Sequence 112, App
c 31	35.4	7.1	558	10	US-09-815-242-133	Sequence 133, App
c 32	35.4	7.1	583	10	US-09-815-242-147	Sequence 147, App
c 33	35.4	7.1	14557	10	US-09-070-927A-367	Sequence 367, App
34	35	7.0	1638	9	US-09-976-740-46	Sequence 46, Appl
35	35	7.0	1638	12	US-10-023-529-46	Sequence 46, Appl
36	35	7.0	1638	12	US-10-023-523-46	Sequence 46, Appl
37	35	7.0	4697	9	US-09-976-740-17	Sequence 17, Appl
38	35	7.0	4697	10	US-09-962-055-17	Sequence 17, Appl
39	35	7.0	4697	12	US-10-023-529-17	Sequence 17, Appl
40	35	7.0	4697	12	US-10-023-523-17	Sequence 17, Appl
c 41	33.6	6.7	891	10	US-09-815-242-9579	Sequence 9579, Ap
42	32.8	6.5	88421	9	US-09-976-059-1	Sequence 1, Appli
c 43	32.2	6.4	2133	9	US-10-124-800-17	Sequence 17, Appl
44	32.2	6.4	2930	10	US-09-745-763-198	Sequence 198, App
c 45	32.2	6.4	8730	9	US-10-124-800-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-736-969A-1
; Sequence 1, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-4 Transmembrane Protein
; FILE REFERENCE: 020054-000411US
; CURRENT APPLICATION NUMBER: US/09/736,969A
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543

```
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;    LENGTH: 6454
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: full length human CLASP-4 cDNA
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (95)..(6121)
;    OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

  Query Match             54.8%;  Score 274.4;  DB 10;  Length 6454;
  Best Local Similarity   71.8%;  Pred. No. 2.6e-74;
  Matches  359;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db    4442 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 4501

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db    4502 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 4561

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           || || | ||| || |||||| | | ||| |||   ||   |||  || | || || ||
Db    4562 AATAATTTTGCAAATAGTGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAA 4621

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||||| ||||| || || ||||| || || |||||||| |||||  |||||||||||||||
Db    4622 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 4681

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db    4682 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 4741

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db    4742 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 4801

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
            | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db    4802 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 4861

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           || |||    | |||    | ||||| | || |||| |   ||||| || || || ||
Db    4862 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 4921

Qy     482 GAGGAGGCCTCCATGATGGA 501
           || || |   | ||||  ||
Db    4922 GAAGAAGGAGCAATGAAAGA 4941


RESULT 2
US-09-736-969A-7
; Sequence 7, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-4 Transmembrane Protein
;  FILE REFERENCE: 020054-000411US
;  CURRENT APPLICATION NUMBER: US/09/736,969A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;    LENGTH: 4391
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: human CLASP-4 cDNA
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (414)..(4058)
;    OTHER INFORMATION: human CLASP-4
US-09-736-969A-7

  Query Match             53.8%;  Score 269.6;  DB 10;  Length 4391;
  Best Local Similarity   71.2%;  Pred. No. 6.6e-73;
  Matches  356;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           || ||   || |||  ||||| |  |||||||| | ||  | || ||  |||| |||||
Db    2379 TATACCAAAAGGAAAACCTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAA 2438

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || |     |  || | | | ||||| ||| ||||  | | ||| |||
Db    2439 CTGATAGCTGATGTAGCACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATC 2498

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           || || | ||| || |||||| | | |||   |   |    |||  || | || || ||
Db    2499 AATAATTTTGCAAATAGTGACAGACCTATGTTGGCAAGAGCCTTTCCCGCAGAAGTCAAA 2558

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||||| ||||| || || ||||| || || |||||||| |||||||||||||||||||||
Db    2559 GACTTGACCAAGAGAATCCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAG 2618

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || ||||| || |||||  | || |||||||| ||| | ||||| |||||||| |||||
Db    2619 AAAGACCCTGAAATGCTAATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACC 2678

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||| || ||||| || |||||||||| ||| ||||| |||||||| |||
Db    2679 CCAGAGCTCAGGAAAACCTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGAT 2738

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
            | |||||||| || ||||| |||||||| ||| |||| ||||| ||||| | |||
Db    2739 TTTTCAGAGGCTGCGATGTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCAT 2798

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
```

```
          || |||    | |||    | |||||  | || |||| |   ||||| || || || ||
Db   2799 CGAAAAAAATTATTTCCTAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGAT 2858

Qy    482 GAGGAGGCCTCCATGATGGA 501
          || || |   | ||||  ||
Db   2859 GAAGAAGGAGCAATGAAAGA 2878


RESULT 3
US-09-864-761-13960
; Sequence 13960, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13960
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011739.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
US-09-864-761-13960

  Query Match           26.1%;  Score 130.6;  DB 10;  Length 483;
  Best Local Similarity  78.1%;  Pred. No. 1.2e-30;
  Matches  157;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy    171 CTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGA 230
          | || ||||||||| | || ||  | ||| | || ||  | |||||||| || |||||||
Db    269 CAGAGGTGAAGGACCTGACTAAGCGTATAAGGACTGTTTTGATGGCCACAGCTCAGATGA 328

Qy    231 AGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCT 290
          ||||||| ||||| ||||| |||||||||||||| |||||||||||||||| || ||||
Db    329 AGGAGCACGAGAAGGACCCCGAGATGCTGGTGGATCTCCAGTACAGCCTGGCAAACTCCT 388

Qy    291 ATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCA 350
          | || ||||| || || ||  | | ||| ||||| || || ||||||| ||| |||| ||
Db    389 ACGCAAGCACTCCTGAACTACGCAGGACCTGGCTGGAAAGTATGGCCAAGATTCATGCCA 448

Qy    351 AAAATGGCGATCTCTCAGAGG 371
          ||| || ||| | || ||||
Db    449 GAAACGGAGATTTATCTGAGG 469


RESULT 4
US-09-736-968A-59
; Sequence 59, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-7 Transmembrane Protein
;  FILE REFERENCE: 020054-000611US
;  CURRENT APPLICATION NUMBER: US/09/736,968A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
;   LENGTH: 2149
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: preliminary human CLASP-7 cDNA sequence
;   NAME/KEY: CDS
;   LOCATION: (2)..(1933)
;   OTHER INFORMATION: human CLASP-7
US-09-736-968A-59
```

```
  Query Match             23.2%;  Score 116.4;  DB 9;  Length 2149;
  Best Local Similarity   55.4%;  Pred. No. 5.6e-26;
  Matches  225;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy      19 CTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGCAGACGTTGT 78
           ||||| |||     |  |||| |||| ||| ||| ||      ||| | |
Db     283 CTTTGCCCGTGTGAAGATGCAGGTCACCATGTCTCTCTCGTCCCTGGTGGGGACGACGCA 342

Qy      79 TGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAG 138
             | |  | |||       | |  |  || ||    | ||||  || || |||  |
Db     343 GAACTTCAGTGAAGAGCACCTGCGACGTTCACTCAAAACCATCCTCACCTATGCTGAGGA 402

Qy     139 TGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGAT 198
           |||  |    |   | ||| || |||| |     | ||  ||||| | |     |   |
Db     403 GGACATGGGGCTGCGGGACAGCACCTTCGCAGAGCAGGTCCAGGACCTGATGTTCAACCT 462

Qy     199 ACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCT 258
           | || | | || | || ||| |   |||||||||| ||  || | ||||| ||||||||
Db     463 GCACATGATCCTGACGGACACGGTGAAGATGAAGGAACACCAGGAGGACCCTGAGATGCT 522

Qy     259 GGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGAC 318
             | ||||||  ||||||  | |||     |||     || | || || ||  ||  |||
Db     523 CATCGACCTCATGTACAGAATTGCCCGGGGCTACCAGGGCTCACCGGACCTTCGGCTGAC 582

Qy     319 GTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAAT 378
           |||  |  | | |||||||  |||  || |   |    ||| | | | | ||||| ||
Db     583 CTGGTTGCAGAACATGGCCGGGAAGCACGCGGAGCTGGGCAACCACGCCGAGGCCGCCCA 642

Qy     379 GTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGG 424
           ||||   || ||||   | ||||| ||||| || || ||| | | |
Db     643 GTGCATGGTGCACGCGGCCGCCCTCGTGGCTGAGTACCTCGCCCTG 688


RESULT 5
US-09-736-968A-1
; Sequence 1, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-7 Transmembrane Protein
;  FILE REFERENCE: 020054-000611US
;  CURRENT APPLICATION NUMBER: US/09/736,968A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;  LENGTH: 6372
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: full length human CLASP-7 cDNA
;  NAME/KEY: CDS
;  LOCATION: (13)..(6156)
;  OTHER INFORMATION: human CLASP-7
US-09-736-968A-1

  Query Match             23.2%;  Score 116.4;  DB 9;  Length 6372;
  Best Local Similarity   55.4%;  Pred. No. 9.2e-26;
  Matches  225;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy      19 CTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGCAGACGTTGT 78
           ||||| |||     |  |||| |||| ||| ||| ||      ||| | |
Db    4506 CTTTGCCCGTGTGAAGATGCAGGTCACCATGTCTCTCTCGTCCCTGGTGGGGACGACGCA 4565

Qy      79 TGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAG 138
             | |  | |||       | |  |  || ||    | ||||  || || |||  |
Db    4566 GAACTTCAGTGAAGAGCACCTGCGACGTTCACTCAAAACCATCCTCACCTATGCTGAGGA 4625

Qy     139 TGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGAT 198
           |||  |    |   | ||| || |||| |     | ||  ||||| | |     |   |
Db    4626 GGACATGGGGCTGCGGGACAGCACCTTCGCAGAGCAGGTCCAGGACCTGATGTTCAACCT 4685

Qy     199 ACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCT 258
           | || | | || | || ||| |   |||||||||| ||  || | ||||| ||||||||
Db    4686 GCACATGATCCTGACGGACACGGTGAAGATGAAGGAACACCAGGAGGACCCTGAGATGCT 4745

Qy     259 GGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGAC 318
             | ||||||  ||||||  | |||     |||     || | || || ||  ||  |||
Db    4746 CATCGACCTCATGTACAGAATTGCCCGGGGCTACCAGGGCTCACCGGACCTTCGGCTGAC 4805

Qy     319 GTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAAT 378
           |||  |  | | |||||||  |||  || |   |    ||| | | | | ||||| ||
Db    4806 CTGGTTGCAGAACATGGCCGGGAAGCACGCGGAGCTGGGCAACCACGCCGAGGCCGCCCA 4865

Qy     379 GTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGG 424
           ||||   || ||||   | ||||| ||||| || || ||| | | |
Db    4866 GTGCATGGTGCACGCGGCCGCCCTCGTGGCTGAGTACCTCGCCCTG 4911


RESULT 6
US-09-864-761-30524
; Sequence 30524, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30524
;   LENGTH: 212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011739.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
;   OTHER INFORMATION: EST_HUMAN HIT: BE896788.1, EVALUE 9.00e-93
;   OTHER INFORMATION: SWISSPROT HIT: O24696, EVALUE 3.90e-01
;   OTHER INFORMATION: NT HIT: gi11429300, EVALUE 6.00e-93
US-09-864-761-30524

  Query Match             23.0%;  Score 115.4;  DB 10;  Length 212;
  Best Local Similarity   79.2%;  Pred. No. 4e-26;
  Matches  137;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy     199 ACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCT 258
           | | || ||  | ||||||| || ||||||||||||| ||||| ||||| ||||||||
Db       1 AAGGACTGTTTTGATGGCCACAGCTCAGATGAAGGAGCACGAGAAGGACCCCGAGATGCT 60

Qy     259 GGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGAC 318
           |||||| |||||||||||||||| || ||||| || ||||| || || ||  | | |||
Db      61 GGTGGATCTCCAGTACAGCCTGGCAAACTCCTACGCAAGCACTCCTGAACTACGCAGGAC 120

Qy     319 GTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGG 371
           |||||  || || ||||||| ||| |||| || ||| || ||| | || ||||
Db     121 CTGGCTGGAAAGTATGGCCAAGATTCATGCCAGAAACGGAGATTTATCTGAGG 173


RESULT 7
US-09-736-960-3
; Sequence 3, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 4026
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: preliminary human CLASP-5 cDNA sequence
;   NAME/KEY: CDS
;   LOCATION: (1)..(3066)
;   OTHER INFORMATION: human CLASP-5
US-09-736-960-3

  Query Match             16.6%;  Score 83.2;  DB 10;  Length 4026;
  Best Local Similarity   50.5%;  Pred. No. 1.2e-15;
  Matches  202;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy      20 TTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGCAGACGTTGTT 79
           ||||   |   | |  ||||||| | ||| ||  |         || | | |   |
Db    1300 TTTGCAAGAGTAAAGATGCAAGTAACCATGTCCCTGGCATCTTTGGTGGGAAGAGCACCA 1359

Qy      80 GGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGT 139
           | | ||   |||        |       ||| ||   |  ||      || | |  |
Db    1360 GACTTTAATGAAGAGCACCTGAGAAGATCCTTGAGGACAATTTTGGCCTATTCAGAAGAG 1419

Qy     140 GACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATA 199
           |||      ||  ||   ||    ||  || |  | ||| ||||  |   |   |  |
Db    1420 GACACAGCCATGCAGATGACTCCTTTTCCCACCCAGGTGGAGGAACTTCTCTGTAATCTG 1479

Qy     200 CGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTG 259
             |   |  ||   | ||| |   | |||| |||   | || | || || ||||||||
Db    1480 AATAGCATCTTATATGACACAGTGAAAATGAGGGAATTTCAGGAAGATCCTGAGATGCTT 1539

Qy     260 GTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACG 319
           |||| |||  ||||||  | |||||    ||        | || || ||  ||  |||
Db    1540 ATGGATCTCATGTACAGAATTGCCAAGAGTTACCAGGCATCTCCTGATCTGCGGCTGACC 1599

Qy     320 TGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATG 379
           |||||| | | ||||||   ||  ||   ||| ||         | | ||||| || |||
Db    1600 TGGCTCCAGAACATGGCAGAGAAACACACCAAGAAGAAGTGCTACACGGAGGCTGCCATG 1659

Qy     380 TGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCA 419
           |||   || ||||   | ||  ||||||| || ||||| |
Db    1660 TGCCTGGTGCACGCCGCTGCGTTAGTGGCTGAGTATCTGA 1699
```

```
RESULT 8
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 7215
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-5 cDNA
;   NAME/KEY: CDS
;   LOCATION: (112)..(6159)
;   OTHER INFORMATION: human CLASP-5
US-09-736-960-1

  Query Match             16.6%;  Score 83.2;  DB 10;  Length 7215;
  Best Local Similarity   50.5%;  Pred. No. 1.5e-15;
  Matches  202;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy      20 TTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGCAGACGTTGTT 79
           ||||   |   | |  ||||||| | ||| ||  |        || | | |   |
Db    4489 TTTGCAAGAGTAAAGATGCAAGTAACCATGTCCCTGGCATCTTTGGTGGGAAGAGCACCA 4548

Qy      80 GGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGT 139
           | | ||   |||       |       ||| ||   |  ||      || | | |   |
Db    4549 GACTTTAATGAAGAGCACCTGAGAAGATCCTTGAGGACAATTTTGGCCTATTCAGAAGAG 4608

Qy     140 GACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATA 199
           |||      ||  ||   ||    ||  || | | ||| ||||  |   |   |   |
Db    4609 GACACAGCCATGCAGATGACTCCTTTTCCCACCCAGGTGGAGGAACTTCTCTGTAATCTG 4668

Qy     200 CGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTG 259
           |   |  ||   | ||| |   | |||| |||   | || | || || ||||||||
Db    4669 AATAGCATCTTATATGACACAGTGAAAATGAGGGAATTTCAGGAAGATCCTGAGATGCTT 4728

Qy     260 GTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACG 319
           ||||  |||  ||||||  | |||||    ||         | || || ||  ||  |||
Db    4729 ATGGATCTCATGTACAGAATTGCCAAGAGTTACCAGGCATCTCCTGATCTGCGGCTGACC 4788

Qy     320 TGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATG 379
           ||||||  | | ||||||   ||  ||   ||| ||         | | |||||  || |||
Db    4789 TGGCTCCAGAACATGGCAGAGAAACACACCAAGAAGAAGTGCTACACGGAGGCTGCCATG 4848

Qy     380 TGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCA 419
           |||   || ||||   | ||  ||||||| || ||||| |
Db    4849 TGCCTGGTGCACGCCGCTGCGTTAGTGGCTGAGTATCTGA 4888


RESULT 9
US-09-796-692-5465
; Sequence 5465, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5465
;   LENGTH: 427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-5465

  Query Match             15.8%;  Score 79.4;  DB 9;  Length 427;
  Best Local Similarity   53.4%;  Pred. No. 6.3e-15;
  Matches  167;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy     107 TCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTC 166
           ||| ||   |  ||      || | |  |   |||      ||  ||   ||    ||
Db      10 TCCTTGAGGACAATTTTGGCCTATTCAGAAGAGGACACAGCCATGCAGATGACTCCTTTT 69

Qy     167 TCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAG 226
           || |  | ||| ||||  |   |   |   |   |  ||   | ||| |   |
Db      70 CCCACCCAGGTGGAGGAACTTCTCTGTAATCTGAATAGCATCTTATATGACACAGTGAAA 129
```

```
Qy     227 ATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAA 286
           |||| |||   | || | || || ||||||||  |||| |||  ||||||  | |||||
Db     130 ATGAGGGAATTTCAGGAAGATCCTGAGATGCTTATGGATCTCATGTACAGAATTGCCAAG 189

Qy     287 TCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCAT 346
             ||        | || || ||  ||  ||| |||||| | | ||||||   ||  ||
Db     190 AGTTACCAGGCATCTCCTGATCTGCGGCTGACCTGGCTCCAGAACATGGCAGAGAAACAC 249

Qy     347 GTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTG 406
           |||  ||         | | ||||| || ||||||   || ||||   | ||  |||||
Db     250 ACCAAGAAGAAGTGCTACACGGAGGCTGCCATGTGCCTGGTGCACGCCGCTGCGTTAGTG 309

Qy     407 GCAGAATATCTCA 419
           || || ||||| |
Db     310 GCTGAGTATCTGA 322


RESULT 10
US-09-796-692-3710
; Sequence 3710, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3710
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (371)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (404)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (411)
;   OTHER INFORMATION: n=A,T,C or G
US-09-796-692-3710

  Query Match           13.8%;  Score 69;  DB 9;  Length 444;
  Best Local Similarity   53.4%;  Pred. No. 1e-11;
  Matches  167;  Conservative    0;  Mismatches  145;  Indels    1;  Gaps    1;

Qy     107 TCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTC 166
           ||| ||   |  ||      || | |  |    |||      ||  ||   ||    ||
Db      10 TCCTTGAGGACAATTTTGGCCTATTCAGAAGAGGACACAGCCATGCAGATGACTCCTTTT 69

Qy     167 TCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAG 226
           || |  | ||| ||||  |   |   |   |    |   |  ||   | ||| |   |
Db      70 CCCACCCAGGTGGAGGAACTTCTCTGTAATCTGAATAGCATCTTATATGACACAGTGAAA 129

Qy     227 ATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAA 286
           |||| |||   | || | || || ||||||||  |||| |||  ||||||  | |||||
Db     130 ATGAGGGAATTTCAGGAAGATCCTGAGATGCTTATGGATCTCATGTACAGAATTGCCAAG 189

Qy     287 TCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCAT 346
             ||        | || || ||  ||  ||| |||||| | | ||||| |||    ||
Db     190 AGTTACCAGGCATCTCCTGATCTGCGGCTGACCTGGCTCCAGAACATGG-CAGAGAACAC 248

Qy     347 GTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTG 406
           |||  ||         | | ||||| || ||||||   || ||||   | ||  |||||
Db     249 ACCAAGAAGAAGTGCTACACGGAGGCTGCCATGTGCCTGGTGCACGCCGCTGCGTTAGTG 308

Qy     407 GCAGAATATCTCA 419
           || || ||||| |
Db     309 GCTGAGTATCTGA 321


RESULT 11
US-09-736-968A-98
; Sequence 98, Application US/09736968A
; Patent No. US20020169283A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-7 Transmembrane Protein
;  FILE REFERENCE: 020054-000611US
;  CURRENT APPLICATION NUMBER: US/09/736,968A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 98
;   LENGTH: 271
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: exon 116900-117170
US-09-736-968A-98

  Query Match             11.5%;  Score 57.4;  DB 9;  Length 271;
  Best Local Similarity   56.7%;  Pred. No. 3e-08;
  Matches  106;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy    238 TGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAG 297
          ||||  |   ||  | || |      ||||  | | |||  | |||     |||     |
Db      4 TGAGTCCCCTCCTCACATCCCCCCTCACCTGGACTCCAGAATTGCCCGGGGCTACCAGGG 63

Qy    298 CACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGG 357
          | | || || ||  ||  ||| ||| |  | | ||||||| |||  || |   |    ||
Db     64 CTCACCGGACCTTCGGCTGACCTGGTTGCAGAACATGGCCGGGAAGCACGCGGACGTGGG 123

Qy    358 CGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCT 417
          | | | | | ||||| ||   ||||   || ||||   | ||||| ||||| || || ||
Db    124 CAACCACGCCGAGGCCGCCCAGTGCATGGTGCACGCGGCCGCCCTCGTGGCTGAGTACCT 183

Qy    418 CACACGG 424
          | | | |
Db    184 CGCCCTG 190


RESULT 12
US-09-864-761-7634
; Sequence 7634, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7634
;   LENGTH: 512
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011472.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
US-09-864-761-7634

  Query Match             11.5%;  Score 57.4;  DB 10;  Length 512;
  Best Local Similarity   56.7%;  Pred. No. 4e-08;
  Matches  106;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy    238 TGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAG 297
          ||||  |   ||  | || |      ||||  | | |||  | |||     |||     |
Db    169 TGAGTCCCCTCCTCACATCCCCCCTCACCTGGACTCCAGAATTGCCCGGGGCTACCAGGG 228

Qy    298 CACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGG 357
          | | || || ||  ||  ||| ||| |  | | ||||||| |||  || |   |    ||
Db    229 CTCACCGGACCTTCGGCTGACCTGGTTGCAGAACATGGCCGGGAAGCACGCGGACGTGGG 288

Qy    358 CGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCT 417
          | | | | | ||||| ||   ||||   || ||||   | ||||| ||||| || || ||
Db    289 CAACCACGCCGAGGCCGCCCAGTGCATGGTGCACGCGGCCGCCCTCGTGGCTGAGTACCT 348

Qy    418 CACACGG 424
          | | | |
Db    349 CGCCCTG 355


RESULT 13
US-09-864-761-24339
; Sequence 24339, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24339
;   LENGTH: 194
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011472.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
;   OTHER INFORMATION: NT HIT: AB037816.1, EVALUE 2.00e-58
;   OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 3.40e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE531136.1, EVALUE 2.00e-87
US-09-864-761-24339

 Query Match             10.0%;  Score 50;  DB 10;  Length 194;
 Best Local Similarity   58.9%;  Pred. No. 4.8e-06;
 Matches   86;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy     279 TGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCA 338
           | |||     |||     || | || || ||  ||  ||| ||| |  | | |||||||
Db       3 TTGCCCGGGGCTACCAGGGCTCACCGGACCTTCGGCTGACCTGGTTGCAGAACATGGCCG 62

Qy     339 GGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAG 398
           |||  || |   |    ||| | | | | ||||| ||   ||||   || ||||   | |
Db      63 GGAAGCACGCGGACGTGGGCAACCACGCCGAGGCCGCCCAGTGCATGGTGCACGCGGCCG 122

Qy     399 CCCTAGTGGCAGAATATCTCACACGG 424
           |||| ||||| || || ||| | | |
Db     123 CCCTCGTGGCTGAGTACCTCGCCCTG 148


RESULT 14
US-09-736-960-76
; Sequence 76, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: 13th exon
US-09-736-960-76

  Query Match             9.7%;  Score 48.8;  DB 10;  Length 266;
  Best Local Similarity   58.1%;  Pred. No. 1.3e-05;
  Matches   86;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     272 TACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGC 331
           | |||  | |||||    ||        | || || ||  ||  ||| |||||| | | |
Db      31 TTCAGAATTGCCAAGAGTTACCAGGCATCTCCTGATCTGCGGCTGACCTGGCTCCAGAAC 90

Qy     332 ATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCAC 391
           |||||   ||  ||   ||| ||         | | ||||| || ||||||   || |||
Db      91 ATGGCAGAGAAACACACCAAGAAGAAGTGCTACACGGAGGCTGCCATGTGCCTGGTGCAC 150

Qy     392 GTAACAGCCCTAGTGGCAGAATATCTCA 419
           |   | ||  ||||||| || ||||| |
Db     151 GCCGCTGCGTTAGTGGCTGAGTATCTGA 178


RESULT 15
US-09-736-960-86
; Sequence 86, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
```

```
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
;   LENGTH: 66686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: ordered human genomic DNA at CLASP-5 locus
US-09-736-960-86

  Query Match             9.7%;  Score 48.8;  DB 10;  Length 66686;
  Best Local Similarity   58.1%;  Pred. No. 0.00016;
  Matches   86;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     272 TACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGC 331
           | |||  | |||||    ||        | || || ||  ||  ||| |||||| | | |
Db   33658 TTCAGAATTGCCAAGAGTTACCAGGCATCTCCTGATCTGCGGCTGACCTGGCTCCAGAAC 33717

Qy     332 ATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCAC 391
           |||||   ||  ||   ||| ||         | | ||||| || ||||||   || |||
Db   33718 ATGGCAGAGAAACACACCAAGAAGAAGTGCTACACGGAGGCTGCCATGTGCCTGGTGCAC 33777

Qy     392 GTAACAGCCCTAGTGGCAGAATATCTCA 419
           |   | ||  |||||| || ||||| |
Db   33778 GCCGCTGCGTTAGTGGCTGAGTATCTGA 33805


Search completed: February  7, 2003, 09:14:41
Job time : 69.3549 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 06:53:49 ; Search time 1010.34 Seconds
(without alignments)
8030.908 Million cell updates/sec

Title: US-09-687-837-1_COPY_2500_3000
Perfect score: 501
Sequence: 1 ttacactggaaagaagtcct.....gaggaggcctccatgatgga 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	499.4	99.7	869	14	BQ434370	BQ434370 AGENCOURT
	2	488.8	97.6	994	14	BQ067940	BQ067940 AGENCOURT
c	3	417.4	83.3	454	10	BE163028	BE163028 QV3-HT045
	4	410.2	81.9	632	12	BG777371	BG777371 602664635
	5	401.4	80.1	694	13	BI691316	BI691316 603309631
	6	370	73.9	889	13	BI695006	BI695006 603345307
	7	341.6	68.2	558	13	BJ070645	BJ070645 BJ070645
	8	337.8	67.4	658	10	BB627919	BB627919 BB627919
	9	330	65.9	503	10	AW909665	AW909665 ur72d11.y
	10	303.8	60.6	520	10	BE533728	BE533728 601232874
	11	295	58.9	922	14	BQ424615	BQ424615 AGENCOURT
c	12	294.8	58.8	871	13	BI249092	BI249092 602992423
	13	291.4	58.2	589	13	BI006821	BI006821 QV3-RT006
	14	276.8	55.2	738	14	BQ179803	BQ179803 UI-M-EW0-
	15	268.6	53.6	1042	13	BM456770	BM456770 AGENCOURT
	16	263	52.5	615	12	BG087170	BG087170 H3136D04-
	17	249.6	49.8	554	10	AW231163	AW231163 uo40c09.y
	18	243.4	48.6	600	12	BG806790	BG806790 2041-42 M
	19	240.4	48.0	733	13	BG919189	BG919189 602818013
	20	232	46.3	346	14	D58813	D58813 HUM512E07B
c	21	230.6	46.0	283	10	AW607903	AW607903 RC0-HT050
	22	228.6	45.6	493	12	BG000343	BG000343 RC3-GN004
c	23	216	43.1	224	10	AW837402	AW837402 QV2-LT003
	24	209.8	41.9	545	13	BI535492	BI535492 399059 MA
	25	209	41.7	595	10	AW211391	AW211391 uo80d10.y
	26	207	41.3	1042	9	AL544595	AL544595 AL544595
	27	206.8	41.3	349	9	AA384900	AA384900 EST98484
	28	204.8	40.9	875	12	BE896788	BE896788 601437471
	29	202	40.3	672	10	AW959926	AW959926 EST371997
	30	195.2	39.0	327	9	AA384941	AA384941 EST98518
	31	194.6	38.8	742	10	BE382198	BE382198 601271772
	32	191.2	38.2	201	12	BF875977	BF875977 RC3-ET013
	33	183.4	36.6	948	14	BQ213756	BQ213756 AGENCOURT
	34	182.6	36.4	314	9	AA327686	AA327686 EST31082
	35	173	34.5	870	14	BQ735599	BQ735599 AGENCOURT
c	36	172.4	34.4	731	10	AV383524	AV383524 AV383524
c	37	171.4	34.2	838	17	CNS00Z0E	AL097064 Drosophil
	38	142.6	28.5	1101	17	CNS050F6	AL315483 Tetraodon
	39	140.4	28.0	652	12	BF979009	BF979009 602147696
	40	136.6	27.3	622	14	BQ390482	BQ390482 NISC_mq13
	41	136.6	27.3	773	12	BG664949	BG664949 DRABTD07
	42	133	26.5	744	13	BG915748	BG915748 602814273
	43	128.6	25.7	466	12	BG345674	BG345674 dd90f03.y
	44	125	25.0	555	12	BF075158	BF075158 223976 MA
	45	121	24.2	966	14	BM905401	BM905401 AGENCOURT

ALIGNMENTS

RESULT 1
BQ434370
LOCUS BQ434370 869 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7827049 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6152826 5', mRNA sequence.
ACCESSION BQ434370
VERSION BQ434370.1 GI:21173446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13491 row: e column: 19
High quality sequence stop: 631.
FEATURES Location/Qualifiers
source 1. .869

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6152826"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      223 a    232 c    236 g    178 t
ORIGIN

  Query Match             99.7%;  Score 499.4;  DB 14;  Length 869;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-130;
  Matches  500;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 176

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     177 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 236

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     237 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 296

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     297 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 356

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 416

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     417 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 476

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     477 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 536

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     537 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 596

Qy     481 CGAGGAGGCCTCCATGATGGA 501
           |||||||||||||||||||||
Db     597 CGAGGAGGCCTCCATGATGGA 617


RESULT 2
BQ067940
LOCUS       BQ067940                994 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AGENCOURT_6792185 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5768931
            5', mRNA sequence.
ACCESSION   BQ067940
VERSION     BQ067940.1  GI:19896986
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 994)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12829  row: j  column: 04
            High quality sequence stop: 499.
FEATURES             Location/Qualifiers
     source          1. .994
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5768931"
                     /clone_lib="NIH_MGC_121"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
                     fetal brains, female age 20 weeks, female age 24 weeks,
                     and male age 26 weeks.  Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     0.7-3.5 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 017. Note:
                     this is a NIH_MGC Library."
BASE COUNT      274 a    255 c    244 g    220 t      1 others
ORIGIN

  Query Match             97.6%;  Score 488.8;  DB 14;  Length 994;
  Best Local Similarity   98.4%;  Pred. No. 1.3e-127;
  Matches  493;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 TTACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCA 246

Qy      61 GCTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCAT 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     247 GCTGATAGCAGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCAT 306

Qy     121 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 CAACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAA 366

Qy     181 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 GGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGA 426

Qy     241 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 GAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCAC 486

Qy     301 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 GCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGA 546

Qy     361 TCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 420
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     547 TCTCTCAGAAGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCAC 606

Qy     421 ACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
           || |||||||||||||||||| |||||||||||||| |||||||||||| || |||||
Db     607 ACNGAAAGGCGTGTTTAGACAAAGATGCACCGCCTTCAAGGTCATTACCCCCAAAATCGA 666

Qy     481 CGAGGAGGCCTCCATGATGGA 501
           |||||| ||||||||||||||
Db     667 CGAGGAAGCCTCCATGATGGA 687


RESULT 3
BE163028/c
LOCUS       BE163028                454 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  QV3-HT0457-070300-113-c03 HT0457 Homo sapiens cDNA, mRNA sequence.
```

```
ACCESSION   BE163028
VERSION     BE163028.1  GI:8625749
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-HT0457-070
            300-113-c03&t3=2000-03-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 3
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0457"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       88 a    112 c    124 g    130 t
ORIGIN

  Query Match             83.3%;  Score 417.4;  DB 10;  Length 454;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-107;
  Matches  418;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 AAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGC 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 AAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGC 395

Qy      70 AGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTG 129
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     394 AGACGTTGTTGGCATTGGGGGAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTG 335

Qy     130 TGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAAC 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 TGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAAC 275

Qy     190 CAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCC 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCC 215

Qy     250 AGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 AGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCT 155

Qy     310 CAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 CAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGA 95

Qy     370 GGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAG 428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAG 36


RESULT 4
BG777371
LOCUS       BG777371                632 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602664635F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4804592 5',
            mRNA sequence.
ACCESSION   BG777371
VERSION     BG777371.1  GI:14047688
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 632)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1651  row: e  column: 09
            High quality sequence stop: 620.
FEATURES             Location/Qualifiers
     source          1. .632
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4804592"
                     /clone_lib="NIH_MGC_60"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      173 a    175 c    167 g    117 t
ORIGIN

  Query Match             81.9%;  Score 410.2;  DB 12;  Length 632;
  Best Local Similarity   99.3%;  Pred. No. 1.9e-105;
  Matches  412;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      87 GGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAGTGACCGGC 146
           ||| ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1 GGGGAACCAGATTCCAGCAGTCCCCCTCCATCATCAACAACTGTGCCAACAGTGACCGGC 60

Qy     147 TTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGG 206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGATACGCACGG 120

Qy     207 TGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACC 266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     121 TGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACC 180

Qy     267 TCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCG 326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCG 240

Qy     327 ACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATG 386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATG 300

Qy     387 TCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAGGCGTGTTTAGACAAGGAT 446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAGGCGTGTTTAGACAAGGAT 360

Qy     447 GCACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGAGGCCTCCATGATGGA 501
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGAGGCCTCCATGATGGA 415
```

RESULT 5
BI691316
LOCUS BI691316 694 bp mRNA linear EST 18-SEP-2001
DEFINITION 603309631F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5345838 5', mRNA sequence.
ACCESSION BI691316
VERSION BI691316.1 GI:15653945
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 694)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11878 row: m column: 07
High quality sequence stop: 693.

FEATURES Location/Qualifiers
source 1. .694
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5345838"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 174 a 182 c 191 g 146 t 1 others
ORIGIN

Query Match 80.1%; Score 401.4; DB 13; Length 694;
Best Local Similarity 87.6%; Pred. No. 6.2e-103;
Matches 438; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||||| |||||||||||| | ||| ||||| || || || |||||||| ||||||||||
Db     146 TACACAGGAAAGAAGTCTTNTGTTCGGACGCACTTACAGGTCATCATCTCTGTCAGCCAA 205

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| ||||| || |||||||||| | |||||||||||||||||| |||| ||||||
Db     206 CTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 265

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||||| |||||| | || |||||||||||||| ||||||||||||||
Db     266 AACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAA 325

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || || || ||||| ||||| || || |||||||| |||||||||||||||| |||
Db     326 GATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAG 385

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||| |||||||||||||||||||||||||||||||| || ||||| ||||||||
Db     386 AACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACC 445

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||||||||||| ||||| ||||| ||||| || || |||||||| ||
Db     446 CCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAATGGGGAC 505

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           |||||||||| |||||||||||||||||||| |||||| | ||||||||||||||||||
Db     506 CTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAGAATATCTCACA 565

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           ||||||||| |||| ||||| || ||||| |||||||||| || || ||||||||||
Db     566 CGGAAAGGCATGTTCAGACAGGGGTGCACAGCCTTCAGGGTTATCACACCAAACATCGAT 625

Qy     482 GAGGAGGCCTCCATGATGGA 501
           || ||||| ||||||||||
Db     626 GAAGAGGCTTCCATGATGGA 645
```

RESULT 6
BI695006
LOCUS BI695006 889 bp mRNA linear EST 18-SEP-2001
DEFINITION 603345307F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5373121 5', mRNA sequence.
ACCESSION BI695006
VERSION BI695006.1 GI:15657635
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11949 row: n column: 02
High quality sequence stop: 652.

FEATURES Location/Qualifiers
source 1. .889
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5373121"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

```
BASE COUNT      253 a    226 c    221 g    189 t
ORIGIN

  Query Match             73.9%;  Score 370;  DB 13;  Length 889;
  Best Local Similarity   87.5%;  Pred. No. 5.7e-94;
  Matches  439;  Conservative    0;  Mismatches   60;  Indels    3;  Gaps    3;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||||| |||||||||| |||||| ||||| || || || |||||||| ||||||||||
Db       8 TACACAGGAAAGAAGT-CTTTGTTCGGACGCACTTACAGGTCATCATCTCTGTCAGCCAA 66

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| ||||| || |||||||||||| | ||||||||||||||||||| |||| ||||||
Db      67 CTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 126

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||||| |||||| | || |||||||||||||| ||||||||||||||
Db     127 AACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAA 186

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || || ||||| ||||| || || |||||||| ||||||||||||||| |||
Db     187 GATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAG 246

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||| |||||||||||||||||||||||||||||| || ||||| ||||||||
Db     247 AACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACC 306

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||||||||||| ||||| ||||| ||||| || || |||||||| ||
Db     307 CCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAATGGGGAC 366

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           |||||||||| |||||||||||||||||||| |||||| | ||||||||||||||||||
Db     367 CTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAGAATATCTCACA 426

Qy     422 CGGAAAGGC-GTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGA 480
           |||||||||  |||| ||||| || ||||| |||||||||| || || ||||||||||
Db     427 CGGAAAGGCACTGTTCAGACAGGGGTGCACAGCCTTCAGGGTTATCACACCAAACATCGA 486

Qy     481 CG-AGGAGGCCTCCATGATGGA 501
           || | ||||| ||||||||||
Db     487 TGAACGAGGCTTCCATGATGGA 508


RESULT 7
BJ070645
LOCUS       BJ070645                558 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ070645 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL091d16 5', mRNA sequence.
ACCESSION   BJ070645
VERSION     BJ070645.1  GI:17500834
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL091d16"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      173 a    116 c    129 g    140 t
ORIGIN

  Query Match             68.2%;  Score 341.6;  DB 13;  Length 558;
  Best Local Similarity   80.2%;  Pred. No. 4.9e-86;
  Matches  401;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           |||||||||||||| || |||||| | |||||  |||| || || || || || || |||
Db      35 TACACTGGAAAGAAATCTTTTGTCAGAACACACCTGCAGGTTATAATCTCAGTAAGTCAG 94

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           |||||||| || || || || || || |  || || |||||||||||||| || ||||||
Db      95 CTGATAGCTGATGTGGTCGGTATAGGTGGCACAAGGTTCCAGCAGTCCCTCTCAATCATC 154

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           ||||| |||||||| || ||| |||| |||||||||||   ||||||||||||||| ||
Db     155 AACAATTGTGCCAATAGCGACAGGCTCATTAAGCACACATCCTTCTCCTCTGATGTTAAA 214

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           ||  | || |||||||| || || ||  | ||||| || || || ||||| || ||||||
Db     215 GATCTGACAAAAAGGATCCGTACCGTCTTGATGGCTACGGCTCAAATGAAAGAACATGAG 274

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           || |||||||| ||| |||||||||| ||||| || || ||||| || ||||| |||||
Db     275 AATGACCCAGAAATGTTGGTGGACCTGCAGTATAGTCTAGCCAAGTCATATGCTAGCACA 334

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || ||||| |||||||| |||||| |||||||||| || |||||||| ||  |||| ||
Db     335 CCTGAGCTGAGGAAGACCTGGCTTGACAGCATGGCAAGAATCCATGTTAAGCATGGTGAC 394

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
           || |||||||||||||||||||||| || ||| |||| | ||  ||||||||||||||
Db     395 CTATCAGAGGCAGCAATGTGCTATGTGCATGTAGCAGCTTTGGTATCAGAATATCTCACA 454

Qy     422 CGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGAC 481
           || |||||  |||| |||||||||||| || | |||||||| || || |||||||| ||
Db     455 CGAAAAGGAATGTTCAGACAAGGATGCTCCACATTCAGGGTGATCACTCCAAACATTGAT 514

Qy     482 GAGGAGGCCTCCATGATGGA 501
           || || || || ||||||||
Db     515 GAAGAAGCTTCTATGATGGA 534


RESULT 8
BB627919
LOCUS       BB627919                658 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB627919 RIKEN full-length enriched, adult male urinary bladder Mus
            musculus cDNA clone 9530048J06 5', mRNA sequence.
ACCESSION   BB627919
VERSION     BB627919.1  GI:16465473
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
```

```
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES                Location/Qualifiers
     source             1. .658
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /clone="9530048J06"
                        /clone_lib="RIKEN full-length enriched, adult male urinary
                        bladder"
                        /sex="male"
                        /tissue_type="urinary bladder"
                        /dev_stage="adult"
                        /lab_host="DH10B"
                        /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                        prepared and sequenced in Mouse Genome Encyclopedia
                        Project of Genome Exploration Research Group in Riken
                        Genomic Sciences Center and Genome Science Laboratory in
                        RIKEN. Division of Experimental Animal Research in Riken
                        contributed to prepare mouse tissues. 1st strand cDNA was
                        primed with a primer [5'
                        GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                        prepared by using trehalose thermo-activated reverse
                        transcriptase and subsequently enriched for full-length by
                        cap-trapper. cDNA went through one round of normalization
                        to Rot = 20.0 and subtraction to Rot = 370.0. Second
                        strand cDNA was prepared with the primer adapter of
                        sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                        3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                        modified pBluescript KS(+) after bulk excision from
                        Lambda FLC I."
BASE COUNT      167 a    176 c    164 g    151 t
ORIGIN

  Query Match             67.4%;  Score 337.8;  DB 10;  Length 658;
  Best Local Similarity   87.8%;  Pred. No. 6.5e-85;
  Matches  380;  Conservative    0;  Mismatches   52;  Indels    1;  Gaps    1;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
                 ||||| |||||||||||| ||||| ||||| || || || |||||||| ||||||||||
Db     226 TACACAGGAAAGAAGTCTTTTGTTCGGACGCACTTACAGGTCATCATCTCTGTCAGCCAA 285

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| ||||| || |||||||||||| | ||||||||||||||||||| |||| ||||||
Db     286 CTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 345

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||||| |||||| | || |||||||||||||| ||||||||||||||
Db     346 AACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAA 405

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || || || ||||| ||||| || || |||||||| |||||||||||||||||| |||
Db     406 GATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAG 465

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||| |||||||||||||||||||||||||||||||| || ||||| ||||||||
Db     466 AACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACC 525

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||||||||||| |||||  ||||| ||||| || || |||||||| ||
Db     526 CCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAATGGGGAC 585

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCC-CTAGTGGCAGAATATCTCAC 420
           |||||||||||  |||||||||||||||||||||  ||||||  | ||||||||||||||||
Db     586 CTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTTGGTGGCAGAATATCTCAC 645

Qy     421 ACGGAAAGGCGTG 433
           |||| ||||| ||
Db     646 ACGGCAAGGCATG 658


RESULT 9
AW909665
LOCUS       AW909665                503 bp    mRNA    linear   EST 25-MAY-2000
DEFINITION  ur72d11.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3155829 5'
            similar to TR:Q63603 Q63603 TRG MRNA ;, mRNA sequence.
ACCESSION   AW909665
VERSION     AW909665.1  GI:8074903
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: ur72d11.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

            MGI:1058585
            Seq primer: -40RP from Gibco
            High quality sequence stop: 416.
FEATURES                Location/Qualifiers
     source             1. .503
                        /organism="Mus musculus"
                        /strain="129,C57BL/6J,FVB/N"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:3155829"
                        /clone_lib="NCI_CGAP_Mam3"
                        /tissue_type="tumor, gross tissue"
                        /dev_stage="10 months"
                        /lab_host="DH10B"
```

```
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
BASE COUNT      130 a    136 c    130 g    107 t
ORIGIN

  Query Match             65.9%;  Score 330;  DB 10;  Length 503;
  Best Local Similarity   87.8%;  Pred. No. 9e-83;
  Matches  360;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||||| |||||||||||| ||||| ||||| || || || |||||||| ||||||||||
Db      93 TACACAGGAAAGAAGTCTTTTGTTCGGACGCACTTACAGGTCATCATCTCTGTCAGCCAA 152

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| ||||| || |||||||||||| | ||||||||||||||||||| |||| ||||||
Db     153 CTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 212

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||||| |||||| | || |||||||||||||| ||||||||||||||
Db     213 AACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAA 272

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || || ||||| ||||| || || |||||||| ||||||||||||||| |||
Db     273 GATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAG 332

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||| |||||||||||||||||||||||||||||| || ||||| ||||||||
Db     333 AACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACC 392

Qy     302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
           || |||||||||||||||||| ||||| ||||| ||||| || || |||||||| ||
Db     393 CCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAATGGGGAC 452

Qy     362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGA 411
           |||||||||||| |||||||||||||||||||| |||||| | ||||||||
Db     453 CTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAGA 502.


RESULT 10
BE533728
LOCUS       BE533728                520 bp    mRNA    linear   EST 09-AUG-2000
DEFINITION  601232874F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3596515 5',
            mRNA sequence.
ACCESSION   BE533728
VERSION     BE533728.1  GI:9762373
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8774  row: d  column: 20
            High quality sequence stop: 520.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3596515"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      136 a    127 c    155 g    102 t
ORIGIN

  Query Match             60.6%;  Score 303.8;  DB 10;  Length 520;
  Best Local Similarity   87.1%;  Pred. No. 2.5e-75;
  Matches  379;  Conservative    0;  Mismatches   52;  Indels    4;  Gaps    4;

Qy      71 GACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAA-CAACTG 129
           || || ||||||| ||| | ||||||||||||||||||| |||| |||||||| ||||||
Db       1 GATGTGGTTGGCAGTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATCAAGCAACTG 60

Qy     130 TGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTT-AA 188
           ||||||||| |||||| | || |||||||||||||| |||||||||||||| || ||  |
Db      61 TGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAAGATTTGGA 120

Qy     189 CCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACC 248
           | || ||||| ||||| || || |||||||| |||||||||||||||| |||||||||
Db     121 CTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAGAACGACC 180

Qy     249 CAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGC 308
           | |||||||||||||||||||||||||||||| || ||||| |||||||| ||||
Db     181 CGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACCCCTGAGC 240

Qy     309 TCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAA-TGGCGATCTCTCA 367
           |||||||||||||||| ||||| ||||| ||||| || || ||||| ||| || ||||||
Db     241 TCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAAGTGGGGACCTCTCA 300

Qy     368 GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAA- 426
           ||||| |||||||||||||||||||| |||||| | ||||||||||||||||||||||
Db     301 GAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAGAATATCTCACACGGAAC 360

Qy     427 AGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGA 486
           |||| |||| ||||| || ||||| |||||||||| || || |||||||||| || ||
Db     361 AGGCATGTTCAGACAGGGGTGCACAGCCTTCAGGGTTATCACACCAAACATCGATGAAGA 420

Qy     487 GGCCTCCATGATGGA 501
           ||| ||||||||||
Db     421 GGCTTCCATGATGGA 435


RESULT 11
BQ424615
LOCUS       BQ424615                922 bp    mRNA    linear   EST 23-MAY-2002
DEFINITION  AGENCOURT_7834017 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6153210
            5', mRNA sequence.
ACCESSION   BQ424615
VERSION     BQ424615.1  GI:21119930
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13492 row: e column: 19
High quality sequence stop: 577.

FEATURES Location/Qualifiers
source 1. .922
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6153210"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 251 a 237 c 249 g 185 t
ORIGIN

Query Match 58.9%; Score 295; DB 14; Length 922;
Best Local Similarity 99.7%; Pred. No. 1.1e-72;
Matches 306; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy    122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AACAACTGTGCC-ACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 59

Qy    182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 119

Qy    242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 179

Qy    302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGAT 239

Qy    362 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACA 299

Qy    422 CGGAAAG 428
          |||||||
Db    300 CGGAAAG 306
```

RESULT 12
BI249092/c
LOCUS BI249092 871 bp mRNA linear EST 17-JUL-2001
DEFINITION 602992423F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5148461 5', mRNA sequence.
ACCESSION BI249092
VERSION BI249092.1 GI:14796120
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 871)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11366 row: m column: 06
High quality sequence stop: 716.

FEATURES Location/Qualifiers
source 1. .871
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5148461"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 175 a 245 c 225 g 226 t
ORIGIN

Query Match 58.8%; Score 294.8; DB 13; Length 871;
Best Local Similarity 80.3%; Pred. No. 1.2e-72;
Matches 408; Conservative 0; Mismatches 92; Indels 8; Gaps 5;

```
Qy      2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCA---AGTCATCATATCTGTCAGC 58
          ||| | |||||||||||| ||||  | |        | |    |||||||| |||||||||
Db    848 TACCCAGGAAAGAAGTCTTTTGGTCAGGACGGCACGTACAGGGTCATCATCTCTGTCAGC 789

Qy     59 CAGC-TGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCAT 117
          || | |||| |||     || || |||||| | ||||||||||||||    ||||  |
Db    788 CAACTTGATTGCAAGATGTGGTGCCATTGGAGGAACCAGATTCCAGCCAGTCCTGGTCTA 729

Qy    118 CATCAACAACTGTGCCAACAGTGACCGGCTT--ATTAAGCACA-CCAGCTTCTCCTCTGA 174
              |||  ||||||||||    | |  |  || ||||||| ||||||| ||||||||
Db    728 TCATCACACTTGTGCCAACAGGCGACCGGATTCATCAAGCACACCCAGCTTTTCCTCTGA 669

Qy    175 TGTGAAGGACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGA 234
          ||||||  || ||  ||  ||  ||||| ||||| ||  ||  |||||||| ||||||||||||
Db    668 TGTGAAAGATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGA 609

Qy    235 GCATGAGAACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGC 294
          |||  ||||||||||  |||||||||||||||||||||||||||||||||  ||  ||||| ||
Db    608 GCACGAGAACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGC 549

Qy    295 CAGCACGCCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAA 354
          ||||||  || ||||||||||||||||||||  |||||  |||||  |||||  ||  ||  |||||
Db    548 CAGCACCCCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCACGTTAAAAA 489

Qy    355 TGGCGATCTCTCAGAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATA 414
          |||  ||  ||||||||||  |||||||||||||||||||||  ||||||  |  ||||||||||
Db    488 TGGGGACCTCTCAGAGGCGGCAATGTGCTATGTCCACGTGACAGCCTTGGTGGCAGAATA 429

Qy    415 TCTCA-CACGGAAAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAA 473
          |||||  ||||||||||  ||||  |||||  ||  |||||  ||||||||||  ||  ||  ||||
Db    428 TCTCACCACGGAAAGGCATGTTCAGACAGGGGTGCACAGCCTTCAGGGTTATCACACCAA 369

Qy    474 ACATCGACGAGGAGGCCTCCATGATGGA 501
          |||||||  ||  |||||  ||||||||||
Db    368 ACATCGATGAAGAGGCTTCCATGATGGA 341
```

RESULT 13
BI006821
LOCUS BI006821 589 bp mRNA linear EST 13-JUN-2001
DEFINITION QV3-RT0067-150101-538-a01 RT0067 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI006821
VERSION BI006821.1 GI:14410879
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-RT0067-
            150101-538-a01&t3=2001-01-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 584.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="RT0067"
                     /dev_stage="Adult"
                     /note="Organ: kidney_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      146 a    164 c    168 g    111 t
ORIGIN

  Query Match             58.2%;  Score 291.4;  DB 13;  Length 589;
  Best Local Similarity   99.7%;  Pred. No. 8.6e-72;
  Matches  292;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     209 CTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTC 268
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17 CCAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCTGGTGGACCTC 76

Qy     269 CAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGAC 328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 CAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGACGTGGCTCGAC 136

Qy     329 AGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTC 388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 AGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAATGTGCTATGTC 196

Qy     389 CACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAGGCGTGTTTAGACAAGGATGC 448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 CACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAGGCGTGTTTAGACAAGGATGC 256

Qy     449 ACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGAGGCCTCCATGATGGA 501
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 ACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGAGGCCTCCATGATGGA 309


RESULT 14
BQ179803
LOCUS       BQ179803                738 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-M-EW0-bwu-d-19-0-UI.r1 NIH_BMAP_EW0 Mus musculus cDNA clone
            IMAGE:5703138 5', mRNA sequence.
ACCESSION   BQ179803
VERSION     BQ179803.1  GI:20355295
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5703138"
                     /clone_lib="NIH_BMAP_EW0"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 15.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is GTGCGTGGAA. This library was created for the
                     University of Iowa Mouse Brain Molecular Anatomy Project
                     (BMAP): 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Instututes of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      182 a    209 c    173 g    174 t
ORIGIN

  Query Match             55.2%;  Score 276.8;  DB 14;  Length 738;
  Best Local Similarity   87.8%;  Pred. No. 1.4e-67;
  Matches  302;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 TACACTGGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAG 61
           ||||| |||||||||| ||||| ||||| || || || |||||||| ||||||||||
Db     395 TACACAGGAAAGAAGTCTTTTGTTCGGACGCACTTACAGGTCATCATCTCTGTCAGCCAA 454

Qy      62 CTGATAGCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATC 121
           ||||| ||||| || |||||||||| | ||||||||||||||||||| |||| ||||||
Db     455 CTGATTGCAGATGTGGTTGGCATTGGAGGAACCAGATTCCAGCAGTCCTTGTCTATCATC 514

Qy     122 AACAACTGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAG 181
           |||||||||||||||| |||||| | || |||||||||||||| ||||||||||||||
Db     515 AACAACTGTGCCAACAGCGACCGGATCATCAAGCACACCAGCTTTTCCTCTGATGTGAAA 574

Qy     182 GACTTAACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAG 241
           || || || || ||||| ||||| || || |||||||| ||||||||||||||||| |||
Db     575 GATTTGACTAAGAGGATCCGCACAGTCCTGATGGCCACAGCCCAGATGAAGGAGCACGAG 634

Qy     242 AACGACCCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACG 301
           |||||||| |||||||||||||||||||||||||||| || ||||| ||||||||
Db     635 AACGACCCGGAGATGCTGGTGGACCTCCAGTACAGCCTGGCTAAGTCCTACGCCAGCACC 694
```

```
Qy    302 CCCGAGCTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCA 345
          || ||||||||||||||||||||| ||||| ||||| ||||| ||
Db    695 CCTGAGCTCAGGAAGACGTGGCTAGACAGTATGGCGAGGATTCA 738
```

RESULT 15
BM456770
LOCUS BM456770 1042 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6404261 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583671 5', mRNA sequence.
ACCESSION BM456770
VERSION BM456770.1 GI:18505810
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12347 row: b column: 24
High quality sequence stop: 685.

FEATURES Location/Qualifiers
source 1. .1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5583671"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 344 a 190 c 221 g 286 t 1 others
ORIGIN

Query Match 53.6%; Score 268.6; DB 13; Length 1042;
Best Local Similarity 72.3%; Pred. No. 3.5e-65;
Matches 349; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```
Qy     19 CTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATAGCAGACGTTGT 78
          |||| |  |||||||| | ||  | || ||  |||| ||||| |||||||| || || |
Db      2 CTTTTTGAGGACACATCTACAGATAATAATTGCTGTAAGCCAACTGATAGCTGATGTAGC 61

Qy     79 TGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAACTGTGCCAACAG 138
             |  || | | | ||||| ||| ||||  | | ||| ||||| || | ||| || ||
Db     62 ACTAAGCGGAGGATCAAGATTTCAGGAGTCTTTATTCATTATCAATAATTTTGCAAATAG 121

Qy    139 TGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTAACCAAAAGGAT 198
          |||| | | ||| |||   ||   |||  || | || || || ||||| ||||| || ||
Db    122 TGACAGACCTATGAAGGCAACTGCCTTTCCCGCAGAAGTCAAAGACTTGACCAAGAGAAT 181

Qy    199 ACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGACCCAGAGATGCT 258
          ||||| || || |||||||| ||||| ||||||||||||||||| ||||| || |||||
Db    182 CCGCACTGTTCTTATGGCCACTGCCCAAATGAAGGAGCATGAGAAAGACCCTGAAATGCT 241

Qy    259 GGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCGAGCTCAGGAAGAC 318
            | || |||||||| ||| | ||||| |||||||| ||||| || ||||||||||| ||
Db    242 AATTGATCTCCAGTATAGCTTAGCCAAGTCCTATGCAAGCACCCCAGAGCTCAGGAAAAC 301

Qy    319 GTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCAGAGGCAGCAAT 378
```

```
          ||||| || |||||||||| ||| ||||| |||||||| ||| | |||||||| || ||
Db    302 CTGGCTTGATAGCATGGCCAAGATTCATGTAAAAAATGGAGATTTTTCAGAGGCTGCGAT 361

Qy    379 GTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAAGGCGTGTTTAG 438
          ||| |||||||| ||| |||| ||||| ||||| | |||    || |||    | |||
Db    362 GTGTTATGTCCATGTAGCAGCTCTAGTTGCAGAGTTTCTTCATCGAAAAAAATTATTTCC 421

Qy    439 ACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATCGACGAGGAGGCCTCCATGAT 498
            | |||||  | || |||| |   ||||| || || || || || || |   | ||||
Db    422 TAACGGATGTTCAGCGTTCAAGAAAATTACTCCCAATATAGATGAAGAAGGAGCAATGAA 481

Qy    499 GGA 501
            ||
Db    482 AGA 484
```

Search completed: February 7, 2003, 08:49:39
Job time : 1018.34 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:39:15 ; Search time 996.663 Seconds
(without alignments)
14629.322 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
Perfect score: 501
Sequence: 1 caatttgtggaagcttgcgg.....ttgtgtcttaacaaaggtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		501	100.0	7506	6	AX255048	AX255048 Sequence
2		490	97.8	3899	6	AX058220	AX058220 Sequence
3		471	94.0	7522	9	AF527605	AF527605 Homo sapi
4		471	94.0	7545	9	AB028981	AB028981 Homo sapi
5	c	365	72.9	155198	2	AL357553	AL357553 Homo sapi
6	c	365	72.9	163316	9	AL161420	AL161420 Human DNA
7		328.4	65.5	2998	10	BC009134	BC009134 Mus muscu
8		278.8	55.6	3227	10	RNTRG	X68101 R.norvegicu
9		220.6	44.0	182640	2	AC126253	AC126253 Mus muscu
10		220.6	44.0	209372	2	AC126033	AC126033 Mus muscu
11		206.2	41.2	175281	2	AC109966	AC109966 Rattus no
12	c	175.6	35.0	192825	2	AC119357	AC119357 Rattus no
13		79.4	15.8	2768	9	AB056820	AB056820 Macaca fa
14		78.2	15.6	2036	9	AK001253	AK001253 Homo sapi
15		78.2	15.6	2413	9	BC015018	BC015018 Homo sapi
16	c	77.6	15.5	168684	2	AC023985	AC023985 Homo sapi
17	c	77.6	15.5	171811	9	AC011739	AC011739 Homo sapi
18	c	74.4	14.9	51808	5	AC104683	AC104683 Danio rer
19		71.2	14.2	162378	2	AC105541	AC105541 Rattus no
20	c	71	14.2	174041	2	AL772340	AL772340 Danio rer
21		67.8	13.5	2422	9	AK000227	AK000227 Homo sapi
22		67.4	13.5	2299	9	AK054649	AK054649 Homo sapi
23		67.4	13.5	2610	6	AX058192	AX058192 Sequence
24		67.4	13.5	3152	9	AK056684	AK056684 Homo sapi
25		67.4	13.5	3472	6	AX173120	AX173120 Sequence
26		67.4	13.5	4391	6	AX173028	AX173028 Sequence
27		67.4	13.5	4393	6	AX173118	AX173118 Sequence
28		67.4	13.5	6454	6	AX173022	AX173022 Sequence
29		65	13.0	153026	9	AL391280	AL391280 Human DNA
30		63.4	12.7	103906	2	AC094791	AC094791 Rattus no
31		63.4	12.7	181127	10	AL672038	AL672038 Mouse DNA
32		56.2	11.2	139887	9	CNS01RGX	AL160233 Human chr
33		44.4	8.9	200100	2	AC115789	AC115789 Mus muscu
34		42.4	8.5	146383	2	AC116367	AC116367 Oryza sat
35		41.8	8.3	244	6	AX174652	AX174652 Sequence
36		41.8	8.3	1476	6	AX399260	AX399260 Sequence
37		41.8	8.3	1605	6	AX174666	AX174666 Sequence
38		41.8	8.3	1669	9	HSM802936	AL583913 Homo sapi
39		41.8	8.3	2014	6	AX174662	AX174662 Sequence
40		41.8	8.3	2904	9	BC019102	BC019102 Homo sapi
41		41.8	8.3	4026	6	AX174571	AX174571 Sequence
42		41.8	8.3	4027	6	AX174664	AX174664 Sequence
43		41.8	8.3	4577	9	AK024436	AK024436 Homo sapi
44		41.8	8.3	5564	9	AK074081	AK074081 Homo sapi
45		41.8	8.3	6422	9	AK090429	AK090429 Homo sapi

ALIGNMENTS

RESULT 1
AX255048
LOCUS AX255048 7506 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 7 from Patent WO0170808.
ACCESSION AX255048
VERSION AX255048.1 GI:16074541
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7506)
AUTHORS Rastelli,L.K. and Gerritsen,M.
TITLE Angiogenesis-associated proteins, and nucleic acids encoding the same

```
  JOURNAL   Patent: WO 0170808-A 7 27-SEP-2001;
            Curagen Corporation (US) ; GENENTECH, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .7506
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     2147 a   1689 c   1764 g   1906 t
ORIGIN

  Query Match           100.0%;  Score 501;  DB 6;  Length 7506;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-133;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6106 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 6165

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6166 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 6225

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6226 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 6285

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6286 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 6345

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6346 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 6405

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6406 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 6465

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6466 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 6525

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6526 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 6585

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
             |||||||||||||||||||||
Db    6586 ATTGTGTCTTAACAAAGGTGT 6606


RESULT 2
AX058220
LOCUS       AX058220                3899 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 90 from Patent WO0077040.
ACCESSION   AX058220
VERSION     AX058220.1  GI:12310721
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3899)
  AUTHORS   Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Yang,J., Reddy,R. and Lu,D.A.
  TITLE     Human intracellular signaling molecules
  JOURNAL   Patent: WO 0077040-A 90 21-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3899
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1078 a    929 c    949 g    943 t
ORIGIN

  Query Match           97.8%;  Score 490;  DB 6;  Length 3899;
  Best Local Similarity  99.8%;  Pred. No. 3.9e-130;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3345 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3404

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3405 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3464

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGA-GAAGACGAGCGTCTTACCGAATTC 179
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    3465 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGACGAAGACGAGCGTCTTACCGAATTC 3524

Qy     180 CCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGAC 239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3525 CCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGAC 3584

Qy     240 CAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTC 299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3585 CAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTC 3644

Qy     300 ATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGA 359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3645 ATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGA 3704

Qy     360 CCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTA 419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3705 CCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTA 3764

Qy     420 TAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTT 479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3765 TAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTT 3824

Qy     480 CATTGTGTCTTAACAAAGGTGT 501
             ||||||||||||||||||||||
Db    3825 CATTGTGTCTTAACAAAGGTGT 3846


RESULT 3
AF527605
LOCUS       AF527605                7522 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens zizimin1 mRNA, complete cds.
ACCESSION   AF527605
VERSION     AF527605.1  GI:22038158
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7522)
  AUTHORS   Meller,N., Irani-Tehrani,M., Kiosses,W.B., Del Pozo,M.A. and
            Schwartz,M.A.
  TITLE     Zizimin1, a novel Cdc42 activator, reveals new guanine nucleotide
            exchange-exchange factor domain for rho proteins
  JOURNAL   Nat. Cell Biol. (2002) In press
REFERENCE   2  (bases 1 to 7522)
  AUTHORS   Meller,N. and Schwartz,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2002) Cell Biology, The Scripps Research
            Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .7522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
     CDS             56. .6265
                     /note="guanine nucleotide exchange factor; Cdc42
                     activator"
                     /codon_start=1
```

```
/product="zizimin1"
/protein_id="AAM90306.1"
/db_xref="GI:22038159"
/translation="MSQPPLLPASAETRKFTRALSKPGTAAELRQSVSEVVRGSVLLA
KPKLIEPLDYENVIVQKKTQILNDCLREMLLFPYDDFQTAILRRQGRYICSTVPAKAE
EEAQSLFVTECIKTYNSDWHLVNYKYEDYSGEFRQLPNKVVKLDKLPVHVYEVDEEVD
KDEDAASLGSQKGGITKHGWLYKGNMNSAISVTMRSFKRRFFHLIQLGDGSYNLNFYK
DEKISKEPKGSIFLDSCMGVVQNNKVRRFAFELKMQDKSSYLLAADSEVEMEEWITIL
NKILQLNFEAAMQEKRNGDSHEDDEQSKLEGSGSGLDSYLPELAKSAREAEIKLKSES
RVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGKRILVKCNDLSFNLQCCVAENEEGPTT
NVEPFFVTLSLFDIKYNRKISADFHVDLNHFSVRQMLATTSPALMNGSGQSPSVLKGI
LHEAAMQYPKQGIFSVTCPHPDIFLVARIEKVLQGSITHCAEPYMKSSDSSKVAQKVL
KNAKQACQRLGQYRMPFAWAARTLFKDASGNLDKNARFSAIYRQDSNKLSNDDMLKLL
ADFRKPEKMAKLPVILGNLDITIDNVSSDFPNYVNSSYIPTKQFETCSKTPITFEVEE
FVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQKSFAKARNIAICIEFKDSDEEDSQPLK
CIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEIKIELPTQLHEKHHLLLTFFHVSCDNS
SKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHIPVSANLPSGYLGYQELGMGRHYG
PEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYCQKTESGAQALGNELVKYLKSL
HAMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVIIHVVAQCHEEGLESHLRSY
VKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNKLLKYSWFFFDVLIKSMA
QHLIENSKVKLLRNQRFPASYHHAVETVVNMLMPHITQKFRDNPEASKNANHSLAVFI
KRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNHEHYIPLNLPMPFG
KGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLREVGTALQEFREVRLIAISVLKNLLIK
HSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMTVKDESLALP
AVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGSLISTDSG
NSLPERNSEKSNSLDKHQQSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALF
TYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARTGMMHARLQQLGSLDNSLTFNHS
YGHSDADVLHQSLLEANIATEVCLTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVY
LCFLQKHQSETALKNVFTALRSLIYKFPSTFYEGRADMCAALCYEILKCCNSKLSSIR
TEASQLLYFLMRNNFDYTGKKSFVRTHLQVIISVSQLIADVVGIGGTRFQQSLSIINN
CANSDRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHENDPEMLVDLQYSLAKSYAST
PELRKTWLDSMARIHVKNGDLSEAAMCYVHVTALVAEYLTRKEAVQWEPPLLPHSHSA
CLRRSRGGVFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVLMELLEQCADGLW
KAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVMHSGRRLLGTYFR
VAFFGQAAQYQFTDSETDVEGFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGS
ENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEM
PFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKK
VAELRQLCSSAEVDMIKLQLKLQGSVSVQVNAGPLAYARAFLDDTNTKRYPDNKVKLL
KEVFRQFVEACGQALAVNERLIKEDQLEYQEEMKANYREMAKELSEIMHEQLG"
BASE COUNT     2146 a    1694 c    1774 g    1908 t
ORIGIN
```

Query Match 94.0%; Score 471; DB 9; Length 7522;
Best Local Similarity 98.2%; Pred. No. 1.2e-124;
Matches 500; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6119 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 6178

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6179 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 6238

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
          |||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db    6239 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 6298

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6299 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 6358

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6359 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 6418

Qy     293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6419 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 6478

Qy     353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6479 GGGAGGACC-AGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 6537
```

```
Qy     413 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6538 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 6597

Qy     473 AAGTTTTCATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||||||||||
Db    6598 AAGTTTTCATTGTGTCTTAACAAAGGTGT 6626
```

```
RESULT 4
AB028981
LOCUS       AB028981                7545 bp    mRNA    linear   PRI 10-MAY-2002
DEFINITION  Homo sapiens mRNA for KIAA1058 protein, partial cds.
ACCESSION   AB028981
VERSION     AB028981.2  GI:20521745
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh12146s1.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
  MEDLINE   99397452
   PUBMED   10470851
REFERENCE   2  (bases 1 to 7545)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
COMMENT     On May 9, 2002 this sequence version replaced gi:5689452.
FEATURES             Location/Qualifiers
     source          1. .7545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="hh12146s1"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
                     /note="This sequence was obtained bysubcloning of the DNA
                     fragments derived from two cDNA clones (1 - 1846 was
                     derived from ef00877 and 1847 - 7545 was derived from
                     hh12146)."
     gene            1. .7545
                     /gene="KIAA1058"
     CDS             <1. .6288
                     /gene="KIAA1058"
                     /codon_start=1
                     /product="KIAA1058 protein"
                     /protein_id="BAA83010.2"
                     /db_xref="GI:20521746"
                     /translation="LACLCLARPGTRSCSSCRLSDLRFQPRMQADKCRTSSRSVKKEL
                     VIESPLQYKDAAQGEVEAESPGPVPAKPKLIEPLDYENVIVQKKTQILNDCLREMLLF
                     PYDDFQTAILRRQGRYICSTVPAKAEEEAQSLFVTECIKTYNSDWHLVNYKYEDYSGE
                     FRQLPNKVVKLDKLPVHVYEVDEEVDKDEDAASLGSQKGGITKHGWLYKGNMNSAISV
                     TMRSFKRRFFHLIQLGDGSYNLNFYKDEKISKEPKGSIFLDSCMGVVQNNKVRRFAFE
                     LKMQDKSSYLLAADSEVEMEEWITILNKILQLNFEAAMQEKRNGDSHEDDEQSKLEGS
                     GSGLDSYLPELAKSAREAEIKLKSESRVKLFYLDPDAQKLDFSSAEPEVKSFEEKFGK
                     RILVKCNDLSFNLQCCVAENEEGPTTNVEPFFVTLSLFDIKYNRKISADFHVDLNHFS
                     VRQMLATTSPALMNGSGQSPSVLKGILHEAAMQYPKQGIFSVTCPHPDIFLVARIEKV
                     LQGSITHCAEPYMKSSDSSKVAQKVLKNAKQACQRLGQYRMPFAWAARTLFKDASGNL
                     DKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSDFPN
                     YVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQK
                     SFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPEEFYDEI
                     KIELPTQLHEKHHLLLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSE
```

```
                     QHIPVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHN
                     FFQYCQKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLTRATQEEVA
                     VNVTRVIIHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKP
                     SADFLTSNKLLKYSWFFFDVLIKSMAQHLIENSKVKLLRNQRFPASYHHAVETVVNML
                     MPHITQKFRDNPEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLF
                     EYKFEFLRVVCNHEHYIPLNLPMPFGKGRIQRYQDLQLDYSLTDEFCRNHFLVGLLLR
                     EVGTALQEFREVRLIAISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQ
                     RINVRDVSPFPVNAGMTVKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASP
                     YTTSTPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSSTLGNSVVRCDK
                     LDQSEIKSLLMCFLYILKSMSDDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRY
                     IARTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSL
                     FTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRSLIYKFPSTFY
                     EGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSFVRTHLQVII
                     SVSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKRIRTVLMATA
                     QMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSEAAMCYVHVT
                     ALVAEYLTRKEAVQWEPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPNIDEEASMME
                     DVGMQDVHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLY
                     DTLHRAYSKVTEVMHSGRRLLGTYFRVAFFGQAAQYQFTDSETDVEGFFEDEDGKEYI
                     YKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPF
                     FDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVK
                     KRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLCSSAEVDMIKLQLKLQGSVSVQVNA
                     GPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGQALAVNERLIKEDQLEYQEE
                     MKANYREMAKELSEIMHEQLG"
BASE COUNT     2175 a   1687 c   1753 g   1930 t
ORIGIN

  Query Match             94.0%;  Score 471;  DB 9;  Length 7545;
  Best Local Similarity   98.2%;  Pred. No. 1.2e-124;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    9;  Gaps    2;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6142 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 6201

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6202 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 6261

Qy     121 GAAATCATGCATGAGCA--------GATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 172
           |||||||||||||||||        ||||||||||||||||||||||||||||||||||
Db    6262 GAAATCATGCATGAGCAGCTGGGATGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTAC 6321

Qy     173 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6322 CGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACG 6381

Qy     233 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6382 GGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTG 6441

Qy     293 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6442 CTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACA 6501

Qy     353 GGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 412
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6502 GGGAGGACC-AGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGA 6560

Qy     413 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6561 CTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCA 6620

Qy     473 AAGTTTTCATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||||||||||
Db    6621 AAGTTTTCATTGTGTCTTAACAAAGGTGT 6649


RESULT 5
AL357553/c
LOCUS       AL357553              155198 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 13 clone RP11-56D6, *** SEQUENCING IN
            PROGRESS ***, 2 unordered pieces.
ACCESSION   AL357553
VERSION     AL357553.8  GI:13620383
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155198)
  AUTHORS   Peck,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 12, 2001 this sequence version replaced gi:12956937.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA56D6
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 152143 bases at least Q40
            Consensus quality: 153069 bases at least Q30
            Consensus quality: 153660 bases at least Q20
            Insert size: 155098; sum-of-contigs
            Insert size: 144134; 11.1% error; agarose-fp
            Quality coverage: 5.59x in Q20 bases; sum-of-contigs Quality
            coverage: 6.57x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2153: contig of 2153 bp in length
            *     2154 2253: gap of      100 bp
            *     2254   155198: contig of 152945 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155198
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-56D6"
                     /clone_lib="RPCI-11.1"
     misc_feature    1. .2153
                     /note="assembly_fragment:00588"
     misc_feature    2254. .155198
                     /note="assembly_fragment:02529
                     clone_end:SP6
                     vector_side:right"
BASE COUNT     43663 a  34926 c  34669 g  41836 t    104 others
ORIGIN

  Query Match             72.9%;  Score 365;  DB 2;  Length 155198;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-94;
  Matches  365;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     137 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130593 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 130534

Qy     197 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130533 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 130474

Qy     257 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 130473 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 130414

Qy     317 GCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGA 376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130413 GCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGA 130354

Qy     377 GAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGG 436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130353 GAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGG 130294

Qy     437 TGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAA 496
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130293 TGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAA 130234

Qy     497 GGTGT 501
           |||||
Db 130233 GGTGT 130229
```

RESULT 6
AL161420/c
LOCUS AL161420 163316 bp DNA linear PRI 01-FEB-2001
DEFINITION Human DNA sequence from clone RP11-155N3 on chromosome 13 Contains ESTs, STSs and GSSs. Contains the 3' part of a novel gene similar to KIAA0694, the KIAA1058 gene and a putative novel gene, complete sequence.
ACCESSION AL161420
VERSION AL161420.10 GI:10443397
KEYWORDS HTG; KIAA0694; KIAA1058.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163316)
AUTHORS Smith,M.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 1, 2000 this sequence version replaced gi:10039689.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr13
This sequence is the entire insert of clone RP11-155N3 The true left end of clone RP11-318G11 is at 114983 in this sequence. The true right end of clone RP11-56D6 is at 42341 in this sequence. The true right end of clone RP11-551M18 is at 43464 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-155N3 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES Location/Qualifiers
source 1. .163316
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-155N3"
/clone_lib="RPCI-11.1"
misc_feature 6. .510
/note="match: GSS: Em:AQ372034"
repeat_region 75. .372
/note="AluJb repeat: matches 1. .298 of consensus"
repeat_region 447. .742
/note="AluSp repeat: matches 1. .296 of consensus"
repeat_region 745. .1025
/note="L1MC1 repeat: matches 6019. .6327 of consensus"
repeat_region 1250. .1566
/note="AluSx repeat: matches 2. .311 of consensus"
repeat_region 1631. .2325
/note="L1MC5 repeat: matches 6683. .7400 of consensus"
repeat_region 2326. .2619
/note="AluSx repeat: matches 1. .295 of consensus"
repeat_region 2620. .2890
/note="L1MC5 repeat: matches 7400. .7665 of consensus"
misc_feature complement(3019. .3311)
/note="match: STS: Em:HSA273ZH1"
misc_feature 3060. .3191
/note="match: GSS: Em:B17811"
misc_feature complement(3065. .3246)
/note="match: STS: Em:AF009212"
misc_feature 3067. .3190
/note="match: GSS: Em:AQ114323"
repeat_region 3087. .3122
/note="18 copies 2 mer tg 97% conserved"
repeat_region 3089. .3124
/note="6 copies 6 mer tgtgtg 97% conserved"
misc_feature 3123. .3280
/note="match: GSS: Em:AQ459714"
repeat_region 3191. .3475
/note="AluSx repeat: matches 1. .288 of consensus"
repeat_region 3692. .3773
/note="MIR repeat: matches 3. .94 of consensus"
repeat_region 3887. .4074
/note="L2 repeat: matches 1607. .1787 of consensus"
repeat_region 4082. .4144
/note="3 copies 21 mer 88% conserved"
repeat_region 4418. .4630
/note="L2 repeat: matches 1354. .1587 of consensus"
repeat_region 7208. .7609
/note="MLT1B repeat: matches 8. .390 of consensus"
repeat_region 9220. .9344
/note="MIR repeat: matches 71. .196 of consensus"
repeat_region 9586. .9667
/note="L1ME repeat: matches 5569. .5651 of consensus"
repeat_region 9714. .9813
/note="L2 repeat: matches 771. .874 of consensus"
repeat_region 9869. .10202
/note="L2 repeat: matches 971. .1261 of consensus"
repeat_region 10203. .10581
/note="MSTA repeat: matches 1. .426 of consensus"
repeat_region 10582. .11267
/note="L2 repeat: matches 1261. .2121 of consensus"
repeat_region 11460. .11570
/note="L2 repeat: matches 2377. .2497 of consensus"
repeat_region 11603. .11778
/note="L1MB3 repeat: matches 5996. .6166 of consensus"
repeat_region 11779. .12160
/note="MER51A repeat: matches 1. .371 of consensus"
repeat_region 12161. .12729
/note="L1MB3 repeat: matches 5403. .5996 of consensus"
repeat_region 12682. .12950
/note="L1MD repeat: matches -1. .274 of consensus"
repeat_region 12945. .13209
/note="L1MC5 repeat: matches 7249. .7517 of consensus"
repeat_region 13323. .13413
/note="AluJ/FRAM repeat: matches 221. .311 of consensus"
repeat_region 13711. .13846
/note="AluSq/x repeat: matches 1. .136 of consensus"
misc_feature 13955. .14417

```
                /note="match: GSS: Em:AQ475161"
     misc_feature    complement(14122. .14590)
                /note="match: GSS: Em:AQ349540"
     repeat_region   14334. .14400
                /note="L1MA10 repeat: matches 6204. .6272 of consensus"
     repeat_region   14407. .14693
                /note="AluSx repeat: matches 1. .288 of consensus"
     misc_feature    14584. .15413
                /note="match: GSS: Em:AQ739876"
     repeat_region   14719. .14793
                /note="L1MC/D repeat: matches 5744. .5818 of consensus"
     repeat_region   15013. .15193
                /note="L1MB3 repeat: matches 5993. .6184 of consensus"
     gene            complement(join(16480. .17741,19199. .19206,20108. .20246,
                20382. .20482,23281. .23471,27992. .28202,30654. .30800,
                31633. .31668,32116. .32121,32345. .32458,33162. .33284,
                47391. .47497,48865. .48933,49823. .49880,52300. .52515,
                52636. .52749,54341. .54523,54642. .54779,60487. .60602,
                68322. .68365,68917. .69053,71585. .71632,73040. .73106,
                76401. .76515,78891. .79024,83174. .83515,86012. .86125,
                86363. .86541,90418. .90529,90921. .91004,91224. .91370,
                102891. .102962,103540. .103665,104551. .104620,
                104885. .105000,106035. .106124,106806. .106892,
                107964. .108085,108701. .108824,109525. .109617,
                111151. .111216,111353. .111540,111625. .111704,
                120481. .120603))
                /gene="bA155N3.2"
     mRNA            complement(join(16480. .17741,19199. .19206,20108. .20246,
                20382. .20482,23281. .23471,27992. .28202,30654. .30800,
                31633. .31668,32116. .32121,32345. .32458,33162. .33284,
                47391. .47497,48865. .48933,49823. .49880,52300. .52515,
                52636. .52749,54341. .54523,54642. .54779,60487. .60602,
                68917. .68984,71585. .71632,73040. .73106,76401. .76515,
                78891. .79024,83398. .83515,86012. .86125,86363. .86541,
                90418. .90529,90921. .91004,91224. .91370,102891. .102962,
                103540. .103665,104551. .104620,104885. .105000,
                106035. .106124,106806. .106892,107964. .108085,
                108701. .108824,109525. .109617,111151. .111216,
                111353. .111540,111625. .111704,120481. .>120603))
                /gene="bA155N3.2"
                /product="bA155N3.2.1 (KIAA1058)"
                /note="match: cDNAs: Em:AB028981 Em:U79289
                match: ESTs: Em:W27872 Em:H29845 Em:H29936 Em:AW160928
                Em:AW457794 Em:T62004 Em:R64450 Em:D58813 Em:H10596
                Em:AW380042 Em:AW344638 Em:AW450751 Em:AF188523
                Em:AA151055 Em:T33377 Em:AW466130 Em:AI014329 Em:Z42960
                Em:AI372817 Em:AI372818 Em:AW389763 Em:AW230012
                Em:AL133724 Em:AW162535 Em:AW607903 Em:AA484945 Em:H28862
                Em:Z39079 Em:AI767696 Em:AW576807 Em:AI078244 Em:R78583
                Em:AA705585 Em:AI271453 Em:AW512964 Em:W88550 Em:AI870702
                Em:AL079478 Em:AI693777 Em:AW576805 Em:AI467994
                Em:AI372819 Em:AI732576 Em:AI277278 Em:AI969485
                Em:AI768094 Em:AA779769 Em:AA502043 Em:R27124 Em:AW576813
                Em:AW576816 Em:AW168032 Em:AI365152 Em:AA165512
                Em:AI120344 Em:AI888868 Em:AA165554 Em:R73986 Em:AA523853
                Em:AI041815 Em:AL134605 Em:AW009821 Em:D62808 Em:AI769357
                Em:AW363093 Em:AI733873 Em:W88496 Em:T03019 Em:AW071481
                Em:AI034000"
                /evidence=not_experimental
     polyA_site      complement(16480)
                /gene="bA155N3.2"
     polyA_site      complement(16482)
                /gene="bA155N3.2"
     polyA_site      complement(16483)
                /gene="bA155N3.2"
     polyA_signal    complement(16498. .16503)
                /gene="bA155N3.2"
     misc_feature    16511. .16637
                /note="match: STS: Em:G42960"
     polyA_signal    complement(16523. .16528)
                /gene="bA155N3.2"
     polyA_signal    complement(16570. .16575)
                /gene="bA155N3.2"
```

```
     CDS             complement(join(17741,19199. .19206,20108. .20246,
                     20382. .20482,23281. .23471,27992. .28202,30654. .30800,

  Query Match             72.9%;  Score 365;  DB 9;  Length 163316;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-94;
  Matches  365;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     137 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 196
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17743 AGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACG 17684

Qy     197 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17683 CCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCG 17624

Qy     257 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17623 TGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGAT 17564

Qy     317 GCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGA 376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17563 GCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGA 17504

Qy     377 GAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGG 436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17503 GAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGG 17444

Qy     437 TGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAA 496
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17443 TGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAA 17384

Qy     497 GGTGT 501
           |||||
Db   17383 GGTGT 17379


RESULT 7
BC009134
LOCUS       BC009134                2998 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to hypothetical protein FLJ20220, clone
            MGC:11827 IMAGE:3596515, mRNA, complete cds.
ACCESSION   BC009134
VERSION     BC009134.1  GI:14318664
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2998)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 16 Row: i Column: 2
```

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

```
FEATURES             Location/Qualifiers
     source          1. .2998
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:11827 IMAGE:3596515"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             381. .1634
                     /codon_start=1
                     /product="Similar to hypothetical protein FLJ20220"
                     /protein_id="AAH09134.1"
                     /db_xref="GI:14318665"
                     /translation="MFRQGCTAFRVITPNIDEEASMMEDVGMQDVHFNEDVLMELLEQ
                     CADGLWKAERYELIADIYKLIIPIYEKRRDFERLAHLYDTLHRAYSKVTEVMHSGRRL
                     LGTYFRVAFFGQGFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQD
                     SGKVNPKDLDSKFAYIQVTHVTPFFDEKELQERRTEFERCHNIRRFMFEMPFTQTGKR
                     QGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLC
                     SSAEVDMIKLQLKLQGSVSVQVNAGPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFV
                     EACGQALAVNERLIKEDQLEYQEEMKANYREMAKELSDIMREQICPLEEKTSVLPNSL
                     HIFNAISGTPTSTVVQGLTSSSSVV"
BASE COUNT      835 a    655 c    768 g    740 t
ORIGIN
```

```
  Query Match             65.5%;  Score 328.4;  DB 10;  Length 2998;
  Best Local Similarity   81.8%;  Pred. No. 1.5e-83;
  Matches  404;  Conservative    0;  Mismatches   86;  Indels    4;  Gaps    2;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             |||||||||||||||| |||||||||| || || ||||| ||||| ||||||||||||
Db    1374 CAATTTGTGGAAGCTTGTGGTCAAGCCTTGGCAGTGAACGAGCGTCTCATTAAAGAAGAC 1433

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           || || ||||| |||||||| ||||| |||||||||||||| ||||| |||||||| ||
Db    1434 CAACTGGAGTACCAGGAAGAGATGAAGGCCAACTACAGGGAGATGGCCAAGGAGCTCTCC 1493

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           || ||||||| ||||||||| |||||||||||||||||||| ||||| |||| ||||||
Db    1494 GACATCATGCGTGAGCAGATTTGCCCCCTGGAGGAGAAGACAAGCGTGCTACCAAATTCC 1553

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           || |||||||||||||||||||| ||||| |||||||||||| ||||||| ||| |||||
Db    1554 CTGCACATCTTCAACGCCATCAGCGGGACACCAACAAGCACAGTGGTTCAAGGGTTGACC 1613

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           |||||||| || || |||||||  | ||||||  |||||||||||||| ||||||||||
Db    1614 AGCTCGTCCTCAGTTGTGTGATTTTACCTCATGAACCGTGTGTGGGGACATGCTTTGTCA 1673

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           | |||||||||||||||  ||||| ||| ||||||||| | | || |||||||| || |
Db    1674 TGTGCAAACTCAGGATGACTTCCAGAGCTAATCACTGGTGTGGCCAAGCACAGGAAGAAG 1733

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           | | ||| |  ||     |     |    | ||   | |||||  || ||||  || |||
Db    1734 CCATGGGGAATGGGAGAGAGAAGGAGCCTGGACTGTGATATTTAATAGCAGA--TTTTAT 1791

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           |||||| |   |||||||||||||| ||||||||||||||||||||||||   |||||||
Db    1792 AGGAGTCGGGGGAAGGTGCACATAGTTTTTTAAATCTCACTGGCAATATT--TAGTTTTC 1849

Qy     481 ATTGTGTCTTAACA 494
             |  ||||||||||
Db    1850 CTCATGTCTTAACA 1863
```

RESULT 8
RNTRG

```
LOCUS       RNTRG                   3227 bp    mRNA    linear   ROD 24-FEB-1999
DEFINITION  R.norvegicus trg mRNA.
ACCESSION   X68101
VERSION     X68101.1  GI:550419
KEYWORDS    trg gene.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3227)
  AUTHORS   Pianese,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1992) L. Pianese, Dipartimento di Biologia,
            Patologia Cellulare e Molecolare, Via Pansini, Naples, ITALY
REFERENCE   2  (bases 1 to 3227)
  AUTHORS   Pianese,L., Porcellini,A., Avvedimento,V.E., D'Esposti,F.,
            Feliciello,A., Monticelli,A., Musti,A.M., Tortora,G., Varrone,S.
            and Cocozza,S.
  TITLE     A novel thyroid transcript negatively regulated by tsh
  JOURNAL   Mol. Biol. 13, 75-83 (1994)
FEATURES             Location/Qualifiers
     source          1. .3227
                     /organism="Rattus norvegicus"
                     /strain="Fischer"
                     /db_xref="taxon:10116"
                     /cell_line="FRTL-"
                     /clone_lib="lambdaGT"
     gene            1. .2218
                     /gene="trg"
     CDS             <1. .2218
                     /gene="trg"
                     /codon_start=2
                     /protein_id="CAA48220.1"
                     /db_xref="GI:550420"
                     /db_xref="SPTREMBL:Q63603"
                     /translation="KLSRGHSPLMKKVFDVYLCFLQKHQSEMALKNVFTALRSLIYKF
                     PSTFYEGRADMCASLCYEVLKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKSFVRTH
                     LQVIISLSQLIADVVGIGGTRFQQSLSIINNCANSDRLIKHTSFSSDVKDLTKRIRTV
                     LMATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLDSMARIHVKNGDLSEAAMC
                     YVHVTALVAEYLTRKEADLALQREPPVFPYSHTSCQRKSRGGMFRQGCTAFRVITPNI
                     DEEASMMEDVGMQDVHFNEDVLMELLEQCADGLWKAERLRAGLLTSINSSSPSMKSGG
                     TLETTHLYDTLHRPYSKVTEVITRAAGSWDLLPGGLFGQGFFEDEDGKEYIYKEPKLT
                     PLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKFAYIQVTHVTPFFDEKELQ
                     ERKTEFERCHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMY
                     QHHTDLNPIEVAIDEMSKKVAELHQLCSSAEVDMIKLQLKLQGSVSVQVNAGPLAYAR
                     AFLDDTNTKRYPDNKVKLLKEVFRQFVEACGQALAVNERLIKEDQLEYQEEMKANYRE
                     IRKELSDIIVPRICPGEDKRATKFPAHLQRHQRDTNKHSGSRVDQFILSCVTLPHEPH
                     VGTCFVMCKLRTTFRANHWFCQAQEEAMGNGREKEPWTVIFNSRFYRSWGKVHIFF"
BASE COUNT      905 a    738 c    777 g    807 t
ORIGIN
```

```
  Query Match             55.6%;  Score 278.8;  DB 10;  Length 3227;
  Best Local Similarity   77.7%;  Pred. No. 3.1e-69;
  Matches  390;  Conservative    0;  Mismatches  102;  Indels   10;  Gaps    4;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||| |||||||||| || |||||||| || || ||||||||||||||||||| |||
Db    1772 CAATTCGTGGAAGCTTGTGGCCAAGCCTTGGCAGTGAACGAACGTCTGATTAAAGAGGAC 1831

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||| ||||| |||||||| |||||||||||||||||||| ||   |||||||| ||
Db    1832 CAGCTGGAGTACCAGGAAGAGATGAAAGCCAACTACAGGGAGATCCGGAAGGAGCTCTCC 1891

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           || |||||      |   || |||||      ||||| ||| |||||  |||| ||||||
Db    1892 GACATCATCGTGCCGAGAATTTGCCC-----TGGAGAGGACAAGCGTGCTACCAAATTCC 1946

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           || |||||||||||||||||| ||||| |||||||||||| ||||||| ||| |||||
Db    1947 CTGCACATCTTCAACGCCATCAGCGGGACACCAACAAGCACAGTGGTTCAAGGGTTGACC 2006

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
```

```
          || || || || || |||||| |  | ||||||  ||    |||||||| ||||||||||
Db    2007 AGTTCATCCTCAGTTGTGTGACTTTACCTCATGAACCACACGTGGGGACATGCTTTGTCA 2066

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGA- 359
          | |||||||||||||| |  || || ||| |||||||||     || |||||||| || |
Db    2067 TGTGCAAACTCAGGACGACTTTCAGAGCTAATCACTGGTTTTGCCAAGCACAGGAAGAAG 2126

Qy     360 CCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTA 419
          ||| |||||| |||||||| |   |     | ||   | |||||  || ||||  || ||
Db    2127 CCATGGGGAATGGGAGAGAGAAAGAGCCCTGGACTGTGATATTTAATAGCAGA--TTTTA 2184

Qy     420 TAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTT 479
          ||||||||   ||||||||||||||||||||||||||||||||||||| ||   ||||||
Db    2185 TAGGAGTTGGGGGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATGTT--TAGTTTT 2242

Qy     480 CATTGTGTCTTAACAAAGGTGT 501
          | |  ||||||||||   || |
Db    2243 CCTCATGTCTTAACAGGTGTAT 2264
```

```
RESULT 9
AC126253
LOCUS       AC126253              182640 bp    DNA     linear   HTG 04-JUL-2002
DEFINITION  Mus musculus chromosome UNK clone RP24-119M15, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC126253
VERSION     AC126253.1  GI:21693957
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 182640)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182640)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT
            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------- Project Information ----------
            Center project name: M_BB0119M15
            ----------
            ---------- Summary Statistics ----------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 177371 bases at least Q40
            Consensus quality: 178021 bases at least Q30
            Consensus quality: 178402 bases at least Q20
            Insert size: 174000; agarose-fp
            Insert size: 181640; sum-of-contigs
            Quality coverage: 11.19 in Q20 bases; agarose-fp
            Quality coverage: 9.89 in Q20 bases; sum-of-contigs
            ----------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4276: contig of 4276 bp in length
            *     4277     4376: gap of unknown length
            *     4377     8990: contig of 4614 bp in length
            *     8991     9090: gap of unknown length
            *     9091    14391: contig of 5301 bp in length
            *    14392    14491: gap of unknown length
            *    14492    23041: contig of 8550 bp in length
            *    23042    23141: gap of unknown length
            *    23142    37848: contig of 14707 bp in length
            *    37849    37948: gap of unknown length
            *    37949    52177: contig of 14229 bp in length
            *    52178    52277: gap of unknown length
            *    52278    75893: contig of 23616 bp in length
            *    75894    75993: gap of unknown length
            *    75994   113549: contig of 37556 bp in length
            *   113550   113649: gap of unknown length
            *   113650   177361: contig of 63712 bp in length
            *   177362   177461: gap of unknown length
            *   177462   179201: contig of 1740 bp in length
            *   179202   179301: gap of unknown length
            *   179302   182640: contig of 3339 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182640
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP24-119M15"
     misc_feature    1. .4276
                     /note="assembly_name:Contig10"
     misc_feature    4377. .8990
                     /note="assembly_name:Contig11"
     misc_feature    9091. .14391
                     /note="assembly_name:Contig12"
     misc_feature    14492. .23041
                     /note="assembly_name:Contig13"
     misc_feature    23142. .37848
                     /note="assembly_name:Contig14"
     misc_feature    37949. .52177
                     /note="assembly_name:Contig15"
     misc_feature    52278. .75893
                     /note="assembly_name:Contig16"
     misc_feature    75994. .113549
                     /note="assembly_name:Contig17"
     misc_feature    113650. .177361
                     /note="assembly_name:Contig18"
     misc_feature    177462. .179201
                     /note="assembly_name:Contig8"
     misc_feature    179302. .182640
                     /note="assembly_name:Contig9"
BASE COUNT    49101 a  42136 c  41309 g  49052 t   1042 others
ORIGIN
```

```
  Query Match           44.0%;  Score 220.6;  DB 2;  Length 182640;
  Best Local Similarity  75.4%;  Pred. No. 2.8e-52;
  Matches  301;  Conservative    0;  Mismatches   94;  Indels    4;  Gaps    2;

Qy      96 CAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGA 155
           ||| | | || |    ||      ||   |  || |     |||| ||||||||||||||
Db 162621 CAGAGCATTGCCCTTTGATAACATTGTCTTTCTGTACCTCGAGATTTGCCCCCTGGAGGA 162680

Qy     156 GAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAAC 215
           |||||| |||||  |||| |||||||| |||||||||||||||||||| ||||| |||||
Db 162681 GAAGACAAGCGTGCTACCAAATTCCCTGCACATCTTCAACGCCATCAGCGGGACACCAAC 162740

Qy     216 AAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGC 275
           ||||||| ||||||| ||| ||||||||||||| || || ||||||||  | ||||||
Db 162741 AAGCACAGTGGTTCAAGGGTTGACCAGCTCGTCCTCAGTTGTGTGATTTTACCTCATGAA 162800

Qy     276 CCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCAC 335
           |||||||||||| |||||||||| |||||||||||||  ||||| ||| ||||||
Db 162801 CCGTGTGTGGGGACATGCTTTGTCATGTGCAAACTCAGGATGACTTCCAGAGCTAATCAC 162860

Qy     336 TGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAA 395
           ||| | | || |||||||| || | | | ||| |  ||     |     |    | ||
```

```
Db 162861 TGGTGTGGCCAAGCACAGGAAGAAGCCATGGGGAATGGGAGAGAGAAGGAGCCTGGACTG 162920

Qy    396 CGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAAT 455
          | |||||  || ||||  || |||||||||  |   ||||||||||||| ||||||||||
Db 162921 TGATATTTAATAGCAGA--TTTTATAGGAGTCGGGGGAAGGTGCACATAGTTTTTTAAAT 162978

Qy    456 CTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACA 494
          |||||||||||||||   ||||||| |  |||||||||
Db 162979 CTCACTGGCAATATT--TAGTTTTCCTCATGTCTTAACA 163015


RESULT 10
AC126033
LOCUS       AC126033               209372 bp    DNA     linear   HTG 02-JUL-2002
DEFINITION  Mus musculus chromosome UNK clone RP24-115A22, WORKING DRAFT
            SEQUENCE, 12 unordered pieces.
ACCESSION   AC126033
VERSION     AC126033.1  GI:21672231
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 209372)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209372)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT
            ---------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ------------- Project Information ------------------
            Center project name: M_BB0115A22
            ----------------------------------------------------
            ------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 203478 bases at least Q40
            Consensus quality: 204725 bases at least Q30
            Consensus quality: 205702 bases at least Q20
            Insert size: 161000; agarose-fp
            Insert size: 208272; sum-of-contigs
            Quality coverage: 14.47 in Q20 bases; agarose-fp
            Quality coverage: 9.74 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2828: contig of 2828 bp in length
            *     2829     2928: gap of unknown length
            *     2929     7171: contig of 4243 bp in length
            *     7172     7271: gap of unknown length
            *     7272    12995: contig of 5724 bp in length
            *    12996    13095: gap of unknown length
            *    13096    23730: contig of 10635 bp in length
            *    23731    23830: gap of unknown length
            *    23831    34449: contig of 10619 bp in length
            *    34450    34549: gap of unknown length
            *    34550    48878: contig of 14329 bp in length
            *    48879    48978: gap of unknown length
            *    48979    66344: contig of 17366 bp in length
            *    66345    66444: gap of unknown length
            *    66445    98995: contig of 32551 bp in length
            *    98996    99095: gap of unknown length
            *    99096   116396: contig of 17301 bp in length
            *   116397   116496: gap of unknown length
            *   116497   135389: contig of 18893 bp in length
            *   135390   135489: gap of unknown length
            *   135490   155507: contig of 20018 bp in length
            *   155508   155607: gap of unknown length
            *   155608   209372: contig of 53765 bp in length.
FEATURES             Location/Qualifiers
     source          1. .209372
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP24-115A22"
     misc_feature    1. .2828
                     /note="assembly_name:Contig11"
     misc_feature    2929. .7171
                     /note="assembly_name:Contig12"
     misc_feature    7272. .12995
                     /note="assembly_name:Contig13"
     misc_feature    13096. .23730
                     /note="assembly_name:Contig14"
     misc_feature    23831. .34449
                     /note="assembly_name:Contig15"
     misc_feature    34550. .48878
                     /note="assembly_name:Contig16"
     misc_feature    48979. .66344
                     /note="assembly_name:Contig17"
     misc_feature    66445. .98995
                     /note="assembly_name:Contig18"
     misc_feature    99096. .116396
                     /note="assembly_name:Contig19"
     misc_feature    116497. .135389
                     /note="assembly_name:Contig20"
     misc_feature    135490. .155507
                     /note="assembly_name:Contig21"
     misc_feature    155608. .209372
                     /note="assembly_name:Contig22"
BASE COUNT    61886 a  42489 c  41685 g  62187 t   1125 others
ORIGIN

  Query Match           44.0%;  Score 220.6;  DB 2;  Length 209372;
  Best Local Similarity   75.4%;  Pred. No. 2.8e-52;
  Matches  301;  Conservative    0;  Mismatches   94;  Indels    4;  Gaps    2;

Qy      96 CAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGA 155
           ||| | | || |    ||      ||   |  || |     |||| ||||||||||||||
Db   84255 CAGAGCATTGCCCTTTGATAACATTGTCTTTCTGTACCTCGAGATTTGCCCCCTGGAGGA 84314

Qy     156 GAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTGGGACTCCAAC 215
           |||||| |||||  |||| |||||||| |||||||||||||||||||| ||||| |||||
Db   84315 GAAGACAAGCGTGCTACCAAATTCCCTGCACATCTTCAACGCCATCAGCGGGACACCAAC 84374

Qy     216 AAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGC 275
           ||||||| ||||||| ||| |||||||||||| || || ||||||||  | ||||||
Db   84375 AAGCACAGTGGTTCAAGGGTTGACCAGCTCGTCCTCAGTTGTGTGATTTTACCTCATGAA 84434

Qy     276 CCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCAC 335
           |||||||||||||| |||||||||| |||||||||||||  ||||| ||| ||||||
Db   84435 CCGTGTGTGGGGACATGCTTTGTCATGTGCAAACTCAGGATGACTTCCAGAGCTAATCAC 84494

Qy     336 TGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAA 395
           ||| | | || |||||||| || | | | ||| |  ||      |     |    | ||
Db   84495 TGGTGTGGCCAAGCACAGGAAGAAGCCATGGGGAATGGGAGAGAGAAGGAGCCTGGACTG 84554

Qy     396 CGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAAT 455
             | |||||  || ||||  || |||||||||  |   ||||||||||||| ||||||||||
Db   84555 TGATATTTAATAGCAGA--TTTTATAGGAGTCGGGGGAAGGTGCACATAGTTTTTTAAAT 84612
```

```
Qy    456 CTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACA 494
          ||||||||||||||||   ||||||| |  ||||||||||
Db  84613 CTCACTGGCAATATT--TAGTTTTCCTCATGTCTTAACA 84649


RESULT 11
AC109966
LOCUS       AC109966               175281 bp    DNA     linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-323I5, *** SEQUENCING IN PROGRESS
            ***, 75 unordered pieces.
ACCESSION   AC109966
VERSION     AC109966.3  GI:21738210
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 175281)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 175281)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 12, 2002 this sequence version replaced gi:18847026.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GQLP
                Center clone name: CH230-323I5
            ---------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 105442 bases at least Q40
                Consensus quality: 111492 bases at least Q30
                Consensus quality: 116270 bases at least Q20
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 75 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1290: contig of 1290 bp in length
            *     1291     1390: gap of unknown length
            *     1391     2414: contig of 1024 bp in length
            *     2415     2514: gap of unknown length
            *     2515     3639: contig of 1125 bp in length
            *     3640     3739: gap of unknown length
            *     3740     5063: contig of 1324 bp in length
            *     5064     5163: gap of unknown length
            *     5164     6243: contig of 1080 bp in length
            *     6244     6343: gap of unknown length
            *     6344     7536: contig of 1193 bp in length
            *     7537     7636: gap of unknown length
            *     7637     9071: contig of 1435 bp in length
            *     9072     9171: gap of unknown length
            *     9172    10200: contig of 1029 bp in length
            *    10201    10300: gap of unknown length
            *    10301    11505: contig of 1205 bp in length
            *    11506    11605: gap of unknown length
            *    11606    12680: contig of 1075 bp in length
            *    12681    12780: gap of unknown length
            *    12781    14056: contig of 1276 bp in length
            *    14057    14156: gap of unknown length
            *    14157    15413: contig of 1257 bp in length
            *    15414    15513: gap of unknown length
            *    15514    16580: contig of 1067 bp in length
            *    16581    16680: gap of unknown length
            *    16681    17692: contig of 1012 bp in length
            *    17693    17792: gap of unknown length
            *    17793    19013: contig of 1221 bp in length
            *    19014    19113: gap of unknown length
            *    19114    20418: contig of 1305 bp in length
            *    20419    20518: gap of unknown length
            *    20519    21880: contig of 1362 bp in length
            *    21881    21980: gap of unknown length
            *    21981    22985: contig of 1005 bp in length
            *    22986    23085: gap of unknown length
            *    23086    24156: contig of 1071 bp in length
            *    24157    24256: gap of unknown length
            *    24257    25732: contig of 1476 bp in length
            *    25733    25832: gap of unknown length
            *    25833    27196: contig of 1364 bp in length
            *    27197    27296: gap of unknown length
            *    27297    28615: contig of 1319 bp in length
            *    28616    28715: gap of unknown length
            *    28716    30145: contig of 1430 bp in length
```

```
*    30146    30245: gap of unknown length
*    30246    31418: contig of 1173 bp in length
*    31419    31518: gap of unknown length
*    31519    32896: contig of 1378 bp in length
*    32897    32996: gap of unknown length
*    32997    34142: contig of 1146 bp in length
*    34143    34242: gap of unknown length
*    34243    35262: contig of 1020 bp in length
*    35263    35362: gap of unknown length
*    35363    36924: contig of 1562 bp in length
*    36925    37024: gap of unknown length
*    37025    38670: contig of 1646 bp in length
*    38671    38770: gap of unknown length
*    38771    40612: contig of 1842 bp in length
*    40613    40712: gap of unknown length
*    40713    42724: contig of 2012 bp in length
*    42725    42824: gap of unknown length
*    42825    43829: contig of 1005 bp in length
*    43830    43929: gap of unknown length
*    43930    45818: contig of 1889 bp in length
*    45819    45918: gap of unknown length
*    45919    48183: contig of 2265 bp in length
*    48184    48283: gap of unknown length
*    48284    49887: contig of 1604 bp in length
*    49888    49987: gap of unknown length
*    49988    51568: contig of 1581 bp in length
*    51569    51668: gap of unknown length
*    51669    53251: contig of 1583 bp in length
*    53252    53351: gap of unknown length
*    53352    55636: contig of 2285 bp in length
*    55637    55736: gap of unknown length
*    55737    57159: contig of 1423 bp in length
*    57160    57259: gap of unknown length
*    57260    59519: contig of 2260 bp in length
*    59520    59619: gap of unknown length
*    59620    63392: contig of 3773 bp in length
*    63393    63492: gap of unknown length
*    63493    65130: contig of 1638 bp in length
*    65131    65230: gap of unknown length
*    65231    68145: contig of 2915 bp in length
*    68146    68245: gap of unknown length
*    68246    70070: contig of 1825 bp in length
*    70071    70170: gap of unknown length
*    70171    73494: contig of 3324 bp in length
*    73495    73594: gap of unknown length
*    73595    75991: contig of 2397 bp in length
*    75992    76091: gap of unknown length
*    76092    77867: contig of 1776 bp in length
*    77868    77967: gap of unknown length
*    77968    80318: contig of 2351 bp in length
*    80319    80418: gap of unknown length
*    80419    82211: contig of 1793 bp in length
*    82212    82311: gap of unknown length
*    82312    84412: contig of 2101 bp in length
*    84413    84512: gap of unknown length
*    84513    85982: contig of 1470 bp in length
*    85983    86082: gap of unknown length
*    86083    88348: contig of 2266 bp in length
*    88349    88448: gap of unknown length
*    88449    90746: contig of 2298 bp in length
*    90747    90846: gap of unknown length
*    90847    92600: contig of 1754 bp in length

  Query Match             41.2%;  Score 206.2;  DB 2;  Length 175281;
  Best Local Similarity   74.7%;  Pred. No. 4e-48;
  Matches  286;  Conservative    0;  Mismatches   93;  Indels    4;  Gaps    2;

Qy     119 CTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATT 178
           |||   |  || |     |||| ||||||||||||||||||| |||||  |||| ||||
Db 133646 CTGTCTTTCTGTACCTCAAGATTTGCCCCCTGGAGGAGAAGACAAGCGTGCTACCAAATT 133705

Qy     179 CCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGA 238
           |||| |||||||||||||||||||||||||| |||||||||||| ||||||| ||| |||

Db 133706 CCCTGCACATCTTCAACGCCATCAGTGGGACACCAACAAGCACAGTGGTTCAAGGGTTGA 133765

Qy     239 CCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGT 298
           |||| || || || || |||||| |  | ||||||  ||    |||||||| ||||||||
Db 133766 CCAGTTCATCCTCAGTTGTGTGACTTTACCTCATGAACCACACGTGGGGACATGCTTTGT 133825

Qy     299 CATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGG 358
           ||| |||||||||||||| |  || || ||| |||||||||     || |||||||| ||
Db 133826 CATGTGCAAACTCAGGACGACTTTCAGAGCTAATCACTGGTTTTGCCAAGCACAGGAAGA 133885

Qy     359 ACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCT 418
           | | | ||| |  ||      |   | |    | ||    | |||||  || ||||  || |
Db 133886 AGCCATGGGGAATGGGAGAGAGAAAGAGCCTGGACTGTGATATTTAATAGCAGA--TTTT 133943

Qy     419 ATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTT 478
           ||||||||||   |||||||||||||||||||||||| ||||||||||  ||   |||||
Db 133944 ATAGGAGTTGGGGGAAGGTGCACATATTTTTTTAAATATCACTGGCAATGTT--TAGTTT 134001

Qy     479 TCATTGTGTCTTAACAAAGGTGT 501
           || |  ||||||||||   || |
Db 134002 TCCTCATGTCTTAACAGGTGTAT 134024


RESULT 12
AC119357/c
LOCUS       AC119357              192825 bp    DNA     linear   HTG 18-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-473M19, *** SEQUENCING IN PROGRESS
            ***, 89 unordered pieces.
ACCESSION   AC119357
VERSION     AC119357.3  GI:21746916
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 192825)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
```

```
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 192825)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 192825)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 14, 2002 this sequence version replaced gi:20429771.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GVFP
                Center clone name: CH230-473M19
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 117932 bases at least Q40
                Consensus quality: 124308 bases at least Q30
                Consensus quality: 129203 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 89 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1129: contig of 1129 bp in length
            *     1130     1229: gap of unknown length
            *     1230     2849: contig of 1620 bp in length
            *     2850     2949: gap of unknown length
            *     2950     4125: contig of 1176 bp in length
            *     4126     4225: gap of unknown length
            *     4226     5644: contig of 1419 bp in length
            *     5645     5744: gap of unknown length
            *     5745     6997: contig of 1253 bp in length
            *     6998     7097: gap of unknown length
            *     7098     8757: contig of 1660 bp in length
            *     8758     8857: gap of unknown length
            *     8858     9860: contig of 1003 bp in length
            *     9861     9960: gap of unknown length
            *     9961    11605: contig of 1645 bp in length
            *    11606    11705: gap of unknown length
            *    11706    12751: contig of 1046 bp in length
            *    12752    12851: gap of unknown length
            *    12852    14490: contig of 1639 bp in length
            *    14491    14590: gap of unknown length
            *    14591    16101: contig of 1511 bp in length
            *    16102    16201: gap of unknown length
            *    16202    17628: contig of 1427 bp in length
            *    17629    17728: gap of unknown length
            *    17729    18883: contig of 1155 bp in length
            *    18884    18983: gap of unknown length
            *    18984    20211: contig of 1228 bp in length
            *    20212    20311: gap of unknown length
            *    20312    21353: contig of 1042 bp in length
            *    21354    21453: gap of unknown length
            *    21454    22518: contig of 1065 bp in length
            *    22519    22618: gap of unknown length
            *    22619    24107: contig of 1489 bp in length
            *    24108    24207: gap of unknown length
            *    24208    25603: contig of 1396 bp in length
            *    25604    25703: gap of unknown length
            *    25704    27171: contig of 1468 bp in length
            *    27172    27271: gap of unknown length
            *    27272    28993: contig of 1722 bp in length
            *    28994    29093: gap of unknown length
            *    29094    30122: contig of 1029 bp in length
            *    30123    30222: gap of unknown length
            *    30223    31317: contig of 1095 bp in length
            *    31318    31417: gap of unknown length
            *    31418    33144: contig of 1727 bp in length
            *    33145    33244: gap of unknown length
            *    33245    34431: contig of 1187 bp in length
            *    34432    34531: gap of unknown length
            *    34532    35743: contig of 1212 bp in length
            *    35744    35843: gap of unknown length
            *    35844    37180: contig of 1337 bp in length
            *    37181    37280: gap of unknown length
            *    37281    38596: contig of 1316 bp in length
            *    38597    38696: gap of unknown length
            *    38697    40031: contig of 1335 bp in length
            *    40032    40131: gap of unknown length
            *    40132    41485: contig of 1354 bp in length
            *    41486    41585: gap of unknown length
            *    41586    42744: contig of 1159 bp in length
            *    42745    42844: gap of unknown length
            *    42845    45158: contig of 2314 bp in length
            *    45159    45258: gap of unknown length
            *    45259    46752: contig of 1494 bp in length
            *    46753    46852: gap of unknown length
            *    46853    48029: contig of 1177 bp in length
            *    48030    48129: gap of unknown length
            *    48130    49682: contig of 1553 bp in length
            *    49683    49782: gap of unknown length
            *    49783    51724: contig of 1942 bp in length
            *    51725    51824: gap of unknown length
            *    51825    53137: contig of 1313 bp in length
            *    53138    53237: gap of unknown length
            *    53238    55287: contig of 2050 bp in length
            *    55288    55387: gap of unknown length
            *    55388    57026: contig of 1639 bp in length
            *    57027    57126: gap of unknown length
            *    57127    58750: contig of 1624 bp in length
            *    58751    58850: gap of unknown length
            *    58851    60180: contig of 1330 bp in length
            *    60181    60280: gap of unknown length
            *    60281    61971: contig of 1691 bp in length
            *    61972    62071: gap of unknown length
            *    62072    63893: contig of 1822 bp in length
            *    63894    63993: gap of unknown length
            *    63994    65782: contig of 1789 bp in length
            *    65783    65882: gap of unknown length
            *    65883    67125: contig of 1243 bp in length
            *    67126    67225: gap of unknown length
            *    67226    69485: contig of 2260 bp in length
            *    69486    69585: gap of unknown length
            *    69586    71428: contig of 1843 bp in length
            *    71429    71528: gap of unknown length
            *    71529    73230: contig of 1702 bp in length
            *    73231    73330: gap of unknown length
            *    73331    75374: contig of 2044 bp in length
            *    75375    75474: gap of unknown length
            *    75475    77927: contig of 2453 bp in length
            *    77928    78027: gap of unknown length
            *    78028    79268: contig of 1241 bp in length
            *    79269    79368: gap of unknown length
            *    79369    81963: contig of 2595 bp in length
```

```
            *    81964    82063: gap of unknown length
            *    82064    83940: contig of 1877 bp in length
            *    83941    84040: gap of unknown length
            *    84041    85690: contig of 1650 bp in length
            *    85691    85790: gap of unknown length
            *    85791    88775: contig of 2985 bp in length

  Query Match             35.0%;  Score 175.6;  DB 2;  Length 192825;
  Best Local Similarity   73.5%;  Pred. No. 2.7e-39;
  Matches  238;  Conservative    0;  Mismatches   84;  Indels    2;  Gaps    1;

Qy     119 CTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATT 178
           |||   |  || |     |||| |||||||||||||||||||| |||||  |||| ||||
Db   62393 CTGTCTTTCTGTACCTCAAGATTTGCCCCCTGGAGGAGAAGACAAGCGTGCTACCAAATT 62334

Qy     179 CCCTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGA 238
           |||| |||||||||||||||||||||||||| |||||||||||| ||||||| ||| |||
Db   62333 CCCTGCACATCTTCAACGCCATCAGTGGGACACCAACAAGCACAGTGGTTCAAGGGTTGA 62274

Qy     239 CCAGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGT 298
           |||| || || || || |||||| |  | ||||||  ||    |||||||| ||||||||
Db   62273 CCAGTTCATCCTCAGTTGTGTGACTTTACCTCATGAACCACACGTGGGGACATGCTTTGT 62214

Qy     299 CATTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGG 358
           ||| ||||||||||||| |  || || ||| |||||||||     || |||||||| ||
Db   62213 CATGTGCAAACTCAGGACGACTTTCAGAGCTAATCACTGGTTTTGCCAAGCACAGGAAGA 62154

Qy     359 ACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCT 418
           | | | ||| |  ||     |   | |    | ||   | |||||  || ||||  || |
Db   62153 AGCCATGGGGAATGGGAGAGAGAAAGAGCCTGGACTGTGATATTTAATAGCAGA--TTTT 62096

Qy     419 ATAGGAGTTGTAAGAAGGTGCACA 442
           ||||||| ||   ||||||||||
Db   62095 ATAGGAGATGGGGGAAGGTGCACA 62072


RESULT 13
AB056820
LOCUS       AB056820               2768 bp    mRNA    linear   PRI 14-MAR-2001
DEFINITION  Macaca fascicularis brain cDNA clone:QflA-15589, full insert
            sequence.
ACCESSION   AB056820
VERSION     AB056820.1  GI:13365942
KEYWORDS    fis (full insert sequence); oligo capping.
SOURCE      Macaca fascicularis adult male frontal lobe left cDNA to mRNA,
            clone_lib:macaque brain cDNA library QflA clone:QflA-15589.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2768)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab host:    TOP10
            Vector:      pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:   DraIII (CACTGTGTG)
            R.  Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by Sugano et al.(University of Tokyo, Institute of
            Medical Science). Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1. .2768
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QflA-15589"
                     /sex="male"
                     /tissue_type="frontal lobe left"
                     /clone_lib="macaque brain cDNA library QflA"
                     /dev_stage="adult"
                     /note="Host: TOP10; Vector: pME18S-FL3 (Acc.No. AB009864);
                     R. Site1: DraIII (CACTGTGTG); R. Site2: DraIII
                     (CACCATGTG)"
     CDS             191. .1525
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB39346.1"
                     /db_xref="GI:13365943"
                     /translation="MCFLHIMKTISYETLIAYWQRAPSPEVSDFFSILDVCLQNFRYL
                     GKRNIIRKIAAAFKFVQSTQNNGTLKGSNPSCQTSGLLSQWMHSTSSHEGHKQHRSQT
                     LPIIRGKNALSNSKLLQMLDNTMTSNSNEIDIVHHVDTEANIATEVCLTILDLLSLFT
                     QTHQRQLQQCDCQNSLMKRVFDTYMLFFQVNQSATALKHVFASLRLFVCKFPSAFFQG
                     PADLCGSFCYEVLKCCNHRSRSTQTEASALLYFFMRKNFEFNKQKSIVRSHLQLIKAV
                     SQLIADAGIGGSRFQHSLAITNNFANGDKQMKNSNFPAEVKDLTKRIRTVLMATAQMK
                     EHEKDPEMLVDLQYSLANSYASTPELRRTWLESMAKIHARNGDLSEAAMCYIHIAALI
                     AEYLKRKGMFSMGWPAFLSITPNIKEEGAVQATDNPDNESPVPLREEENTSR"
BASE COUNT      861 a    588 c    570 g    749 t
ORIGIN

  Query Match             15.8%;  Score 79.4;  DB 9;  Length 2768;
  Best Local Similarity   71.7%;  Pred. No. 9.8e-12;
  Matches  104;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           |||||||  || || || || || ||| | |  || || || || || || ||||| |||
Db    1773 CAATTTGCAGATGCATGTGGGCAGGCCCTTGACGTGAATGAGCGCCTCATCAAAGAGGAC 1832

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||| ||||| ||||||||| |||   || |||||| ||| |||   |  || || ||
Db    1833 CAGCTGGAGTACCAGGAAGAACTGAGGTCCCACTACAAGGACATGCTCAGCGAACTCTCC 1892

Qy     121 GAAATCATGCATGAGCAGATCTGCC 145
             |||||| |||||||| |||| |
Db    1893 ACTATCATGAATGAGCAGCTCTGTC 1917


RESULT 14
AK001253
LOCUS       AK001253               2036 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ10391 fis, clone NT2RM4000139, moderately
            similar to R.norvegicus trg mRNA.
ACCESSION   AK001253
VERSION     AK001253.1  GI:7022393
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA,
            clone_lib:NT2RM4 clone:NT2RM4000139.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2036)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .2036
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4000139"
                     /cell_line="NT2"
                     /cell_type="teratocarcinoma"
                     /clone_lib="NT2RM4"
                     /note="cloning vector: pME18SFL3-mRNA from uninduced NT2
                     neuronal precursor cells."
     CDS             153. .1400
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA91583.1"
                     /db_xref="GI:7022394"
                     /translation="MFSMGWPAFLSITPNIKEEGAMKEDSGMQDTPYNENILVEQLYM
                     CVEFLWKSERYELIADVDKPIIAVFEKQRDFKKLSDLYYDIHRSYLKVAEVVNSEKRL
                     FGRYYRVAFYGQGFFEEEEGKEYIYKEPKLTGLSEISQRLLKLYADKFGADNVKIIQD
                     SNKVNPKDLDPKYAYIQVTYVTPFFEEKEIEDRKTDFEMHHNINRFVFETPFTLSGKK
                     HGGVAEQCKRRTILTTSHLFPYVKKRIQVISQSSTELNPIEVAIDEMSKKVSELNQLC
                     TMEEVDMIRLQLKLQGSVSVKVNAGPMAYARAFLEETNAKKYPDNQVKLLKEIFRQFA
                     DACGQALDVNERLIKEDQLEYQEELRSHYKDMLSELSTVMNEQITGRDDLSKRGVDQT
                     CTRVISKATPALPTVSISSSAEV"
BASE COUNT      634 a    390 c    460 g    552 t
ORIGIN

  Query Match             15.6%;  Score 78.2;  DB 9;  Length 2036;
  Best Local Similarity   67.5%;  Pred. No. 2.1e-11;
  Matches  110;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           |||||||  || || || || || ||| | |  || || || || || || ||||| |||
Db    1146 CAATTTGCAGATGCATGTGGGCAGGCCCTTGACGTGAATGAGCGCCTCATCAAAGAGGAC 1205

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||| ||||| ||||||||| |||   || |||||| ||| |||   |  || || ||
Db    1206 CAGCTGGAGTACCAGGAAGAACTGAGGTCCCACTACAAGGACATGCTCAGCGAACTCTCC 1265

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGA 163
             | ||||| ||||||||||      |  ||| ||   | | |
Db    1266 ACAGTCATGAATGAGCAGATTACGGGCAGGGACGACCTGTCAA 1308


RESULT 15
BC015018
LOCUS       BC015018                2413 bp    mRNA    linear   PRI 04-OCT-2001
DEFINITION  Homo sapiens, clone MGC:8871 IMAGE:3922744, mRNA, complete cds.
ACCESSION   BC015018
VERSION     BC015018.1  GI:15929124
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2413)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: p Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 8923209.
FEATURES             Location/Qualifiers
     source          1. .2413
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:8871 IMAGE:3922744"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             139. .1767
                     /codon_start=1
                     /product="Unknown (protein for MGC:8871)"
                     /protein_id="AAH15018.1"
                     /db_xref="GI:15929125"
                     /translation="MKNSNFPAEVKDLTKRIRTVLMATAQMKEHEKDPEMLVDLQYSL
                     ANSYASTPELRRTWLESMAKIHARNGDLSEAAMCYIHIAALIAEYLKRKGYWKVEKIC
                     TASLLSEDTHPCDSNSLLTTPSGGSMFSMGWPAFLSITPNIKEEGAMKEDSGMQDTPY
                     NENILVEQLYMCVEFLWKSERYELIADVNKPIIAVFEKQRDFKKLSDLYYDIHRSYLK
                     VAEVVNSEKRLFGRYYRVAFYGQGFFEEEEGKEYIYKEPKLTGLSEISQRLLKLYADK
                     FGADNVKIIQDSNKVNPKDLDPKYAYIQVTYVTPFFEEKEIEDRKTDFEMHHNINRFV
                     FETPFTLSGKKHGGVAEQCKRRTILTTSHLFPYVKKRIQVISQSSTELNPIEVAIDEM
                     SKKVSELNQLCTMEEVDMIRLQLKLQGSVSVKVNAGPMAYARAFLEETNAKKYPDNQV
                     KLLKEIFRQFADACGQALDVNERLIKEDQLEYQEELRSHYKDMLSELSTVMNEQITGR
                     DDLSKRGVDQTCTRVISKATPALPTVSISSSAEV"
BASE COUNT      761 a    477 c    549 g    626 t
ORIGIN

  Query Match             15.6%;  Score 78.2;  DB 9;  Length 2413;
  Best Local Similarity   67.5%;  Pred. No. 2.1e-11;
  Matches  110;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           |||||||  || || || || || ||| | |  || || || || || || ||||| |||
Db    1513 CAATTTGCAGATGCATGTGGGCAGGCCCTTGACGTGAATGAGCGCCTCATCAAAGAGGAC 1572

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||| ||||| ||||||||| |||   || |||||| ||| |||   |  || || ||
Db    1573 CAGCTGGAGTACCAGGAAGAACTGAGGTCCCACTACAAGGACATGCTCAGCGAACTCTCC 1632

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGA 163
             | ||||| ||||||||||      |  ||| ||   | | |
Db    1633 ACAGTCATGAATGAGCAGATTACGGGCAGGGACGACCTGTCAA 1675


Search completed: February  7, 2003, 08:11:28
Job time : 1465.66 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:36:04 ; Search time 128.085 Seconds
(without alignments)
8808.598 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
Perfect score: 501
Sequence: 1 caatttgtggaagcttgcgg.....ttgtgtcttaacaaaggtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	501	100.0	1493	22	AAI86535	Human polynucleoti
	2	501	100.0	3614	24	ABK84967	DNA encoding cadhe
	3	501	100.0	4806	24	ABK84964	DNA encoding cadhe
	4	501	100.0	4807	21	AAC87972	Human CLASP-2 nucl
	5	501	100.0	4807	21	AAC87973	Human CLASP-2A nuc
	6	501	100.0	4807	24	ABK84966	DNA encoding cadhe
	7	501	100.0	4807	24	ABK84973	DNA encoding cadhe
	8	501	100.0	4898	21	AAC87974	Preliminary CLASP-
	9	501	100.0	4898	21	AAC87975	Preliminary CLASP-
	10	501	100.0	4898	21	AAC87976	Preliminary CLASP-
	11	501	100.0	4898	21	AAC87977	Preliminary CLASP-
	12	501	100.0	4898	21	AAC87978	Preliminary CLASP-
	13	501	100.0	4898	21	AAC87979	Preliminary CLASP-
	14	501	100.0	4898	21	AAC87980	Preliminary CLASP-
	15	501	100.0	4898	21	AAC87981	Preliminary CLASP-
	16	501	100.0	4898	24	ABK84992	DNA encoding cadhe
	17	501	100.0	4898	24	ABK84993	DNA encoding cadhe
	18	501	100.0	4898	24	ABK84994	DNA encoding cadhe
	19	501	100.0	4898	24	ABK84995	DNA encoding cadhe
	20	501	100.0	4898	24	ABK84996	DNA encoding cadhe
	21	501	100.0	4898	24	ABK84997	DNA encoding cadhe
	22	501	100.0	4898	24	ABK84998	DNA encoding cadhe
	23	501	100.0	4898	24	ABK84999	DNA encoding cadhe
	24	501	100.0	6791	24	ABK85003	DNA encoding cadhe
	25	501	100.0	6816	21	AAC74524	Human ORFX ORF79 p
	26	501	100.0	7506	22	AAD19118	Angiogenesis assoc
	27	490	97.8	3899	22	AAF32675	Human cDNA encodin
	28	483	96.4	5048	24	ABK84965	DNA encoding cadhe
	29	471	94.0	5862	24	ABK84970	DNA encoding cadhe
	30	261	52.1	1250	24	ABK84969	DNA encoding cadhe
	31	261	52.1	3642	24	ABK84971	DNA encoding cadhe
	32	261	52.1	3705	24	ABK84968	DNA encoding cadhe
	33	79.4	15.8	2165	21	AAC76949	Human ORFX ORF2504
	34	78.2	15.6	2036	22	AAH14086	Human cDNA sequenc
	35	78.2	15.6	5214	21	AAC87968	Mouse CLASP-1 nucl
	36	78.2	15.6	5214	21	AAA14824	DNA encoding a mur
	37	78.2	15.6	5688	21	AAC87969	Human CLASP relate
	38	78.2	15.6	5688	21	AAA14825	DNA encoding a hum
	39	78.2	15.6	7277	24	AAS18951	Human cDNA encodin
c	40	78	15.6	662	21	AAC87999	Human CLASP-2 geno
c	41	78	15.6	662	24	ABK84991	DNA encoding cadhe
	42	67.4	13.5	2610	22	AAF32647	Human cDNA encodin
	43	67.4	13.5	3472	22	AAS08358	Human cDNA encodin
	44	67.4	13.5	4393	22	AAS08357	Human cDNA encodin
	45	67.4	13.5	6454	22	AAS08334	Human cDNA encodin

ALIGNMENTS

RESULT 1
AAI86535/c
ID AAI86535 standard; cDNA; 1493 BP.
XX
AC AAI86535;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 6595.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO06604.

XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 6595; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1493 BP; 466 A; 314 C; 249 G; 464 T; 0 other;

Query Match 100.0%; Score 501; DB 22; Length 1493;
Best Local Similarity 100.0%; Pred. No. 2.1e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1404 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 1345

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1344 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 1285

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1284 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 1225

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1224 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 1165

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1164 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 1105

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1104 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 1045

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1044 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 985

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     984 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 925

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||
Db     924 ATTGTGTCTTAACAAAGGTGT 904
```

RESULT 2
ABK84967
ID ABK84967 standard; cDNA; 3614 BP.
XX
AC ABK84967;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform #2.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61673.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or

```
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC   PCR primers of the invention.
XX
SQ   Sequence 3614 BP; 1009 A; 834 C; 874 G; 897 T; 0 other;

  Query Match            100.0%;  Score 501;  DB 24;  Length 3614;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2607 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 2666

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2667 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 2726

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2727 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 2786

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2787 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 2846

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2847 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 2906

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2907 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 2966

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2967 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 3026

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3027 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 3086

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    3087 ATTGTGTCTTAACAAAGGTGT 3107


RESULT 3
ABK84964
ID   ABK84964 standard; cDNA; 4806 BP.
XX
AC   ABK84964;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP).
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200231117-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
PA   (ARBO-) ARBOR VITA CORP.
PA   (GARM/) GARMAN J D.
PA   (CAND/) CANDIA A F.
XX
PI   Lu PS;
XX
DR   WPI; 2002-416861/44.
DR   P-PSDB; ABG61670.
XX
PT   New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT   an immune response, and for treating multiple sclerosis, rheumatoid
PT   arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT   and sepsis -
XX
PS   Disclosure; Figure 1; 245pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising an amino
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g., agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. CLASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC   PCR primers of the invention.
XX
SQ   Sequence 4806 BP; 1352 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match            100.0%;  Score 501;  DB 24;  Length 4806;
  Best Local Similarity  100.0%;  Pred. No. 3.8e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3799 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3858

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3859 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3918

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3919 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 3978

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3979 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4038

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4039 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4098

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4099 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4158

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4159 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4218

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4219 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4278

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4279 ATTGTGTCTTAACAAAGGTGT 4299
```

RESULT 4
AAC87972
ID AAC87972 standard; cDNA; 4807 BP.
XX
AC AAC87972;
XX
DT 07-MAR-2001 (first entry)
XX
DE Human CLASP-2 nucleotide sequence.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
DR P-PSDB; AAB36527.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Example 1; Fig 1; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence encodes human CLASP-2, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.8e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3800 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3859

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3860 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3919

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3920 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 3979

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3980 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4039

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4040 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4099

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4100 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4159

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4160 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4219

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4220 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4279

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4280 ATTGTGTCTTAACAAAGGTGT 4300
```

```
RESULT 5
AAC87973
ID   AAC87973 standard; cDNA; 4807 BP.
XX
AC   AAC87973;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Human CLASP-2A nucleotide sequence.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
DR   P-PSDB; AAB36528.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Example 1; Fig 2B; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence encodes human CLASP-2A, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4807;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3800 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3859

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3860 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3919

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3920 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 3979

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3980 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4039

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4040 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4099

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4100 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4159

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4160 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4219

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4220 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4279

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
             |||||||||||||||||||||
Db    4280 ATTGTGTCTTAACAAAGGTGT 4300


RESULT 6
ABK84966
ID   ABK84966 standard; cDNA; 4807 BP.
XX
AC   ABK84966;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding cadherin-like asymmetry protein (CLASP) isoform #1.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200231117-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
```

PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61672.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Disclosure; Figure 3A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity
CC of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC antibodies or are used as therapeutic polypeptides. The CLASP-2
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides can screen for CLASP-2 agonists and antagonists.
CC CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC disorders of the immune system, by activating or inhibiting the
CC activation, differentiation of immune cells and can treat or detect
CC deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC or polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC represent cadherin-like asymmetry protein (CLASP) coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 4807;
Best Local Similarity 100.0%; Pred. No. 3.8e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3800 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3859

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3860 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3919

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3920 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 3979

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3980 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4039

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4040 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4099

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4100 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4159

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4160 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4219

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4220 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4279

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||
Db    4280 ATTGTGTCTTAACAAAGGTGT 4300
```

RESULT 7
ABK84973
ID ABK84973 standard; DNA; 4807 BP.
XX
AC ABK84973;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding cadherin-like asymmetry protein (CLASP) isoform.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation;
KW CLASP; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32202.
XX
PR 13-OCT-2000; 2000US-0687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
DR P-PSDB; ABG61686.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating
PT an immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis -
XX
PS Example 4; Figure 6A; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino

```
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. (II) is useful for
CC   inhibiting a immune response in a subject. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC   polynucleotides can screen for CLASP-2 agonists and antagonists.
CC   CLASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. CLASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g., agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. CLASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (CLASP) coding sequences
CC   and PCR primers of the invention.
XX
SQ   Sequence 4807 BP; 1353 A; 1117 C; 1136 G; 1201 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 24;  Length 4807;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3800 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3859

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3860 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 3919

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3920 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 3979

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3980 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4039

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4040 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4099

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4100 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4159

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4160 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4219

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4220 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4279

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4280 ATTGTGTCTTAACAAAGGTGT 4300


RESULT 8
AAC87974
ID   AAC87974 standard; cDNA; 4898 BP.
XX
AC   AAC87974;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #1.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
DR   P-PSDB; AAB36529.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10A; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
```

CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
             |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

RESULT 9
AAC87975
ID AAC87975 standard; cDNA; 4898 BP.
XX
AC AAC87975;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #2.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10B; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130
```

```
Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

RESULT 10
AAC87976
ID AAC87976 standard; cDNA; 4898 BP.
XX
AC AAC87976;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #3.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10C; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

RESULT 11
AAC87977
ID AAC87977 standard; cDNA; 4898 BP.
XX
AC AAC87977;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #4.

```
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10D; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391


RESULT 12
AAC87978
ID   AAC87978 standard; cDNA; 4898 BP.
XX
AC   AAC87978;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #5.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
```

PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10E; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

```
  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
          |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

RESULT 13
AAC87979
ID AAC87979 standard; cDNA; 4898 BP.
XX
AC AAC87979;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #6.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10F; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.

```
XX
SQ   Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391


RESULT 14
AAC87980
ID   AAC87980 standard; cDNA; 4898 BP.
XX
AC   AAC87980;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Preliminary CLASP-2 nucleotide sequence #7.
XX
KW   CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW   cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW   immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW   hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW   hypersensitivity; transplantation rejection response; immunodeficiency;
KW   proliferation; differentiation; inflammatory response; arthritis;
KW   inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW   anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW   endometriosis; pregnancy induced hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061747-A2.
XX
PD   19-OCT-2000.
XX
PF   13-APR-2000; 2000WO-US10158.
XX
PR   14-APR-1999;   99US-0129171.
PR   14-MAY-1999;   99US-0134114.
PR   14-MAY-1999;   99US-0134117.
PR   14-MAY-1999;   99US-0134118.
PR   21-OCT-1999;   99US-0160860.
PR   29-OCT-1999;   99US-0162498.
PR   13-DEC-1999;   99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Lu PS;
XX
DR   WPI; 2000-619230/59.
XX
PT   Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT   polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT   inflammatory responses -
XX
PS   Disclosure; Fig 10G; 286pp; English.
XX
CC   The present invention describes cadherin-like asymmetry protein-2
CC   (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC   antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC   antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC   used to inhibit an immune response in a subject by interfering with the
CC   ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC   immune response in a subject may also be inhibited by administering an
CC   antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC   proteins and antibodies can be used to prevent or treat a CLASP-2
CC   mediated disease, such as an autoimmune disease caused or exacerbated
CC   by increased activity of TH11 cells. They can also be used to treat
CC   hypersensitivities, prevent transplantation rejection responses and
CC   augment immune responsiveness in immunodeficiency states, inhibit
CC   proliferation and differentiation of cells involved in an inflammatory
CC   response e,g, arthritis, inflammatory bowel disease and increase
CC   differentiation and proliferation of haematopoietic cells e.g. to treat
CC   anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC   treated by disrupting CLASP-2 function include multiple sclerosis,
CC   rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC   The present sequence represents a preliminary CLASP-2 nucleotide
CC   sequence, from the present invention.
XX
SQ   Sequence 4898 BP; 1379 A; 1134 C; 1166 G; 1219 T; 0 other;

  Query Match             100.0%;  Score 501;  DB 21;  Length 4898;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-143;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250
```

```
Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

RESULT 15
AAC87981
ID AAC87981 standard; cDNA; 4898 BP.
XX
AC AAC87981;
XX
DT 07-MAR-2001 (first entry)
XX
DE Preliminary CLASP-2 nucleotide sequence #8.
XX
KW CLASP-1; CLASP-2; transmembrane protein; immune response; inflammatory;
KW cadherin-like asymmetry protein; autoimmune disease; immunosuppressive;
KW immunomodulatory; antiinflammatory; antiarthritic; cytostatic;
KW hypotensive; antirheumatic; antianaemic; haemostatic; neuroprotective;
KW hypersensitivity; transplantation rejection response; immunodeficiency;
KW proliferation; differentiation; inflammatory response; arthritis;
KW inflammatory bowel disease; haematopoietic cell; blood protein disorder;
KW anaemia; thrombocytopaenia; multiple sclerosis; rheumatoid arthritis;
KW endometriosis; pregnancy induced hypertension; ss.
XX
OS Homo sapiens.
XX
PN WO200061747-A2.
XX
PD 19-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10158.
XX
PR 14-APR-1999; 99US-0129171.
PR 14-MAY-1999; 99US-0134114.
PR 14-MAY-1999; 99US-0134117.
PR 14-MAY-1999; 99US-0134118.
PR 21-OCT-1999; 99US-0160860.
PR 29-OCT-1999; 99US-0162498.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu PS;
XX
DR WPI; 2000-619230/59.
XX
PT Isolated cadherin-like asymmetry protein-2 polynucleotide and
PT polypeptide used to diagnose, treat and prevent autoimmune diseases and
PT inflammatory responses -
XX
PS Disclosure; Fig 10H; 286pp; English.
XX
CC The present invention describes cadherin-like asymmetry protein-2
CC (CLASP-2). CLASP-2 can have immunosuppressive, immunomodulatory,
CC antiinflammatory, antiarthritic, cytostatic, hypotensive, antirheumatic,
CC antianaemic, haemostatic and neuroprotective activities. CLASP-2 can be
CC used to inhibit an immune response in a subject by interfering with the
CC ability of a CLASP-2 protein to bind to another T cell or B cell. An
CC immune response in a subject may also be inhibited by administering an
CC antibody which specifically binds to CLASP-2. CLASP-2 polynucleotides,
CC proteins and antibodies can be used to prevent or treat a CLASP-2
CC mediated disease, such as an autoimmune disease caused or exacerbated
CC by increased activity of TH11 cells. They can also be used to treat
CC hypersensitivities, prevent transplantation rejection responses and
CC augment immune responsiveness in immunodeficiency states, inhibit
CC proliferation and differentiation of cells involved in an inflammatory
CC response e,g, arthritis, inflammatory bowel disease and increase
CC differentiation and proliferation of haematopoietic cells e.g. to treat
CC anaemia, thrombocytopaenia and other blood protein disorders. Disorders
CC treated by disrupting CLASP-2 function include multiple sclerosis,
CC rheumatoid arthritis, endometriosis and pregnancy induced hypertension.
CC The present sequence represents a preliminary CLASP-2 nucleotide
CC sequence, from the present invention.
XX
SQ Sequence 4898 BP; 1376 A; 1140 C; 1162 G; 1220 T; 0 other;

Query Match 100.0%; Score 501; DB 21; Length 4898;
Best Local Similarity 100.0%; Pred. No. 3.9e-143;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 3950

Qy      61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3951 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 4010

Qy     121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 4070

Qy     181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 4130

Qy     241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 4190

Qy     301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGAC 4250

Qy     361 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 CAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTAT 4310

Qy     421 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 AGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTTTTC 4370

Qy     481 ATTGTGTCTTAACAAAGGTGT 501
           |||||||||||||||||||||
Db    4371 ATTGTGTCTTAACAAAGGTGT 4391
```

Search completed: February 7, 2003, 07:08:42
Job time : 163.085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:00:34 ; Search time 28.6858 Seconds
(without alignments)
5356.145 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
Perfect score: 501
Sequence: 1 caatttgtggaagcttgcgg.....ttgtgtcttaacaaaggtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		139	27.7	1031	4	US-09-397-787-157	Sequence 157, App
2		41.8	8.3	2225	4	US-09-484-970B-33	Sequence 33, Appl
3	c	40.8	8.1	7218	1	US-08-232-463-14	Sequence 14, Appl
4	c	33	6.6	36519	3	US-08-923-137-2	Sequence 2, Appli
5	c	32.8	6.5	2570	2	US-09-056-075-2	Sequence 2, Appli
6	c	32.8	6.5	2613	4	US-09-255-829-7	Sequence 7, Appli
7	c	32.8	6.5	3891	1	US-08-480-604A-27	Sequence 27, Appl
8	c	32.8	6.5	3891	2	US-08-405-496A-27	Sequence 27, Appl
9	c	32.8	6.5	3891	4	US-08-915-136-27	Sequence 27, Appl
10		32.4	6.5	289	4	US-09-007-005-17	Sequence 17, Appl
11		32.4	6.5	289	4	US-09-244-796-17	Sequence 17, Appl
12	c	32.4	6.5	6559	4	US-09-234-186-1	Sequence 1, Appli
13	c	32.4	6.5	6559	4	US-09-233-527-1	Sequence 1, Appli
14	c	32.4	6.5	6560	5	PCT-US93-05651-1	Sequence 1, Appli
15		32	6.4	1241	1	US-08-471-033-39	Sequence 39, Appl
16		32	6.4	1241	1	US-08-471-033-42	Sequence 42, Appl
17		32	6.4	1241	2	US-08-471-044-39	Sequence 39, Appl
18		32	6.4	1241	2	US-08-471-044-42	Sequence 42, Appl
19		32	6.4	1241	2	US-08-463-483A-39	Sequence 39, Appl
20		32	6.4	1241	2	US-08-463-483A-42	Sequence 42, Appl
21		32	6.4	1241	2	US-08-471-046A-39	Sequence 39, Appl
22		32	6.4	1241	2	US-08-471-046A-42	Sequence 42, Appl
23		32	6.4	1241	2	US-08-470-566B-39	Sequence 39, Appl
24		32	6.4	1241	2	US-08-470-566B-42	Sequence 42, Appl
25		32	6.4	1241	2	US-08-469-334-39	Sequence 39, Appl
26		32	6.4	1241	2	US-08-469-334-42	Sequence 42, Appl
27		32	6.4	1241	3	US-09-300-529-39	Sequence 39, Appl
28		32	6.4	1241	3	US-09-300-529-42	Sequence 42, Appl
29		32	6.4	1358	1	US-08-471-033-45	Sequence 45, Appl
30		32	6.4	1358	2	US-08-471-044-45	Sequence 45, Appl
31		32	6.4	1358	2	US-08-463-483A-45	Sequence 45, Appl
32		32	6.4	1358	2	US-08-471-046A-45	Sequence 45, Appl
33		32	6.4	1358	2	US-08-470-566B-45	Sequence 45, Appl
34		32	6.4	1358	2	US-08-469-334-45	Sequence 45, Appl
35		32	6.4	1358	3	US-09-300-529-45	Sequence 45, Appl
36		32	6.4	1389	1	US-08-471-033-27	Sequence 27, Appl
37		32	6.4	1389	2	US-08-471-044-27	Sequence 27, Appl
38		32	6.4	1389	2	US-08-463-483A-27	Sequence 27, Appl
39		32	6.4	1389	2	US-08-471-046A-27	Sequence 27, Appl
40		32	6.4	1389	2	US-08-470-566B-27	Sequence 27, Appl
41		32	6.4	1389	2	US-08-469-334-27	Sequence 27, Appl
42		32	6.4	1389	3	US-09-300-529-27	Sequence 27, Appl
43		32	6.4	1399	1	US-08-471-033-24	Sequence 24, Appl
44		32	6.4	1399	2	US-08-471-044-24	Sequence 24, Appl
45		32	6.4	1399	2	US-08-463-483A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-397-787-157
; Sequence 157, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 157
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-157

Query Match 27.7%; Score 139; DB 4; Length 1031;
Best Local Similarity 96.6%; Pred. No. 4.3e-33;
Matches 142; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 570

Qy     61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 630

Qy    121 GAAATCATGCATGAGCAGATCTGCCCC 147
          |||||||||||||||||| |  |  ||
Db    631 GAAATCATGCATGAGCAGGTGAGGGCC 657
```

RESULT 2
US-09-484-970B-33
; Sequence 33, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B

```
;  CURRENT FILING DATE:  2000-01-18
;  NUMBER OF SEQ ID NOS:  172
;  SOFTWARE:  PERL Program
; SEQ ID NO 33
;   LENGTH: 2225
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6426186 238544.2CB1
US-09-484-970B-33

  Query Match             8.3%;  Score 41.8;  DB 4;  Length 2225;
  Best Local Similarity   56.0%;  Pred. No. 0.0035;
  Matches   79;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      16 TGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAG 75
           || ||| |||  ||| |  |||| | ||||| || |  | |||||||   || ||||||
Db     721 TGTGGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAGACCAGAGGGAATATCAG 780

Qy      76 GAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAG 135
             | ||| | |||   ||||| |   |  |     || | ||      ||| ||  |   |
Db     781 CAGGAACTCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGCGG 840

Qy     136 CAGATCTGCCCCCTGGAGGAG 156
           | ||       ||| |  ||
Db     841 AAAATTCCAGAACTGTACAAG 861
.

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             8.1%;  Score 40.8;  DB 1;  Length 7218;
  Best Local Similarity   6.8%;  Pred. No. 0.012;
  Matches   30;  Conservative  213;  Mismatches  195;  Indels    0;  Gaps    0;

Qy       8 TGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCG 67
           | ||||  |  |||  ::    :: :: ::: ::: :   ::  ::::::::  ::   :
Db    1450 TAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391

Qy      68 AGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCA 127
           :: :  :::::::: ::::::  ::  : :::::: :: :::::::      ::::  :
Db    1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331

Qy     128 TGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACA 187
           : : ::: :::   :      ::::::::::: ::: :    :  :::        : :
Db    1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271

Qy     188 TCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGT 247
             :: :  :  :: ::::    :: ::: : :: ::   : :::: ::  ::   :
Db    1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211

Qy     248 CTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAA 307
             ::  : : ::  : :    : ::   : : : ::::   :    :  :  : ::
Db    1210 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1151

Qy     308 ACTCAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGG 367
           :   :::: :      ::::  ::  : :::::::  ::: : ::::::::  ::::::
Db    1150 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1091

Qy     368 AAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTT 427
           :::::::::::::::::: ::::::|   |    |  ||  ||   |  | |     |
Db    1090 RRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGAA 1031

Qy     428 GTAAGAAGGTGCACATAT 445
           |||    |||  | |||
Db    1030 TTAATTCTGTGAGCGTAT 1013


RESULT 4
US-08-923-137-2/c
; Sequence 2, Application US/08923137
; Patent No. 6083716
;  GENERAL INFORMATION:
;    APPLICANT:  Wilson, James M.
;    APPLICANT:  Farina, Steven F.
;    APPLICANT:  Fisher, Krishna J.
;    TITLE OF INVENTION:  Chimpanzee Adenovirus Vectors
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr., P.O. Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  United States of America
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/923,137
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,700
;      FILING DATE:  06-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
```

```
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  GNVPN.021CIP1USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36519 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
US-08-923-137-2

  Query Match             6.6%;  Score 33;  DB 3;  Length 36519;
  Best Local Similarity   53.5%;  Pred. No. 6;
  Matches   69;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy      67 GAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATC 126
           ||| |  | || ||| || | |    |    |||  | | |  ||| |   |  | |
Db    9252 GAGGAAGAAGAGGAAGTGGAGGAACTCCCGGGGGCCTTCGAGCGCGAGGTGCGCGCCACC 9193

Qy     127 ATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCAC 186
           ||    ||||  ||| | |  ||||||||| ||  || |||||   ||     |   ||
Db    9192 ATCGCCGAGCTCATCCGTCTTCTGGAGGAGGAGTTGACCGTCTCGGCGCGCAACTCCCAG 9133

Qy     187 ATCTTCAAC 195
           | ||||||
Db    9132 TTTTTCAAC 9124


RESULT 5
US-09-056-075-2/c
; Sequence 2, Application US/09056075
; Patent No. 5955368
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Eric A.
;    APPLICANT:  Bradshaw, Marite
;    APPLICANT:  Rood, Julian
;    TITLE OF INVENTION:  Expression System for Clostridium
;    TITLE OF INVENTION:  Species
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,075
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27386
;      REFERENCE/DOCKET NUMBER:  960296.95238
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2570 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-056-075-2

  Query Match             6.5%;  Score 32.8;  DB 2;  Length 2570;
  Best Local Similarity   58.0%;  Pred. No. 2.1;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     402 TTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACT 461
           ||   ||||  ||| | | |  ||||| |  ||    | || ||  ||||  | ||| ||
Db    1539 TTAACAACATTCTTCCAAGATCAGTTGAATAAATTCTCTCAAATAATTTTGTAACTCCCT 1480

Qy     462 GGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
             |||| | |   |||||||| | | |    |||  |||
Db    1479 TTAAATAATTATCTTTTTCATTATCTGTACTTAAATATGT 1440


RESULT 6
US-09-255-829-7/c
; Sequence 7, Application US/09255829
; Patent No. 6461617
;  GENERAL INFORMATION:
;    APPLICANT:  Shone, Clifford Charles
;    APPLICANT:  Quinn, Conrad Padraig
;    APPLICANT:  Foster, Keith Alan
;    TITLE OF INVENTION:  Recombinant Toxin Fragments
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/255,829
;      FILING DATE:  23-FEB-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB97/02273
;      FILING DATE:  22-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/782,893
;      FILING DATE:  27-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ESMOND, ROBERT W.
;      REGISTRATION NUMBER:  32,893
;      REFERENCE/DOCKET NUMBER:  1581.0130002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2613 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2613
US-09-255-829-7

  Query Match             6.5%;  Score 32.8;  DB 4;  Length 2613;
  Best Local Similarity   58.0%;  Pred. No. 2.1;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     402 TTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACT 461
           ||   ||||  ||| | | |  ||||| |  ||    | || ||  ||||  | ||| ||
Db     325 TTAACAACATTCTTCCAAGATCAGTTGAATAAATTCTCTCAAATAATTTTGTAACTCCCT 266

Qy     462 GGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
```

```
              |||| | |   |||||||| | | |    |||  |||
Db    265 TTAAATAATTATCTTTTTCATTATCTGTACTTAAATATGT 226


RESULT 7
US-08-480-604A-27/c
; Sequence 27, Application US/08480604A
; Patent No. 5736139
;  GENERAL INFORMATION:
;    APPLICANT:  KINK, JOHN A.
;    APPLICANT:  THALLEY, BRUCE S.
;    APPLICANT:  PADHYE, NISHA V.
;    APPLICANT:  FIRCA, JOSEPH R.
;    APPLICANT:  STAFFORD, DOUGLAS C.
;    TITLE OF INVENTION:  VACCINE AND ANTITOXIN FOR TREATMENT AND
;    TITLE OF INVENTION:  PREVENTION OF C. DIFFICILE DISEASE
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
:      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,604A
;      FILING DATE:  07-JUN-1995
:      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/422,711
;      FILING DATE:  14-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/405,496
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01763
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3891 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3888
US-08-480-604A-27

  Query Match             6.5%;  Score 32.8;  DB 1;  Length 3891;
  Best Local Similarity   58.0%;  Pred. No. 2.6;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy    402 TTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACT 461
          ||   ||||  ||| | | |  ||||| |  ||    | || ||  ||||  | ||| ||
Db    325 TTAACAACATTCTTCCAAGATCAGTTGAATAAATTCTCTCAAATAATTTTGTAACTCCCT 266

Qy    462 GGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
             |||| | |   |||||||| | | |    |||  |||
Db    265 TTAAATAATTATCTTTTTCATTATCTGTACTTAAATATGT 226


RESULT 8
US-08-405-496A-27/c
; Sequence 27, Application US/08405496A
; Patent No. 5919665
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAMS, JAMES A.
;    TITLE OF INVENTION:  VACCINE FOR CLOSTRIDIUM BOTULINUM
;    TITLE OF INVENTION:  NEUROTOXIN
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,496A
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3891 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3888
US-08-405-496A-27

  Query Match             6.5%;  Score 32.8;  DB 2;  Length 3891;
  Best Local Similarity   58.0%;  Pred. No. 2.6;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy    402 TTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACT 461
          ||   ||||  ||| | | |  ||||| |  ||    | || ||  ||||  | ||| ||
Db    325 TTAACAACATTCTTCCAAGATCAGTTGAATAAATTCTCTCAAATAATTTTGTAACTCCCT 266
```

```
Qy     462 GGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
           ||||  | |   |||||||| | | |    |||  |||
Db     265 TTAAATAATTATCTTTTTCATTATCTGTACTTAAATATGT 226


RESULT 9
US-08-915-136-27/c
; Sequence 27, Application US/08915136
; Patent No. 6290960
;  GENERAL INFORMATION:
;    APPLICANT:  KINK, JOHN A.
;    APPLICANT:  THALLEY, BRUCE S.
;    APPLICANT:  PADHYE, NISHA V.
;    APPLICANT:  FIRCA, JOSEPH R.
;    APPLICANT:  STAFFORD, DOUGLAS C.
;    TITLE OF INVENTION:  VACCINE AND ANTITOXIN FOR TREATMENT AND
;    TITLE OF INVENTION:  PREVENTION OF C. DIFFICILE DISEASE
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,136
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/480,604
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/405,496
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01763
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3891 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3888
US-08-915-136-27

  Query Match             6.5%;  Score 32.8;  DB 4;  Length 3891;
  Best Local Similarity   58.0%;  Pred. No. 2.6;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     402 TTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACT 461
           ||   ||||  ||| | | |  ||||| |  ||    | || ||  ||||  | ||| ||
Db     325 TTAACAACATTCTTCCAAGATCAGTTGAATAAATTCTCTCAAATAATTTTGTAACTCCCT 266

Qy     462 GGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
           ||||  | |   |||||||| | | |    |||  |||
Db     265 TTAAATAATTATCTTTTTCATTATCTGTACTTAAATATGT 226


RESULT 10
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

  Query Match             6.5%;  Score 32.4;  DB 4;  Length 289;
  Best Local Similarity   5.4%;  Pred. No. 1.1;
  Matches   11;  Conservative   87;  Mismatches  106;  Indels    0;  Gaps    0;

Qy     191 TCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTT 250
           : |:|:  |  :|  : ::    ::: : :::   :   : :::   ::  :  ::
Db      44 URARCRARARURGRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNR 103

Qy     251 CGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACT 310
           ::   : : :   : :       ::  : : : : :::   ::     ::   ::: :
Db     104 SRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNR 163

Qy     311 CAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAG 370
           :: :  :     :: :  ::   ::  : : :   : : : ::: : :   : ::: :|:
Db     164 SRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRCRAR 223

Qy     371 GGGAGAGAAAGGAAATAAAGAACA 394
           |: : :|: :|: :| |: : :|:
Db     224 GRCRURGRCRGRURARARCRURCR 247


RESULT 11
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
```

```
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

  Query Match             6.5%;  Score 32.4;  DB 4;  Length 289;
  Best Local Similarity   5.4%;  Pred. No. 1.1;
  Matches   11;  Conservative   87;  Mismatches  106;  Indels    0;  Gaps    0;

Qy     191 TCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTT 250
           : |:|:  |  :|  : ::    ::: : :::   :   : :::   ::  :  ::
Db      44 URARCRARARURGRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNR 103

Qy     251 CGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCATTTGCAAACT 310
           ::   : : :   : :       ::  : : : : :::   ::     ::   ::: :
Db     104 SRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNR 163

Qy     311 CAGGATGCTTTCCAAAGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAG 370
           :: :  :     :: :  ::   ::  : : :   : : : ::: : :   : ::: :|:
Db     164 SRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRCRAR 223

Qy     371 GGGAGAGAAAGGAAATAAAGAACA 394
           |: : :|: :|: :| |: : :|:
Db     224 GRCRURGRCRGRURARARCRURCR 247


RESULT 12
US-09-234-186-1/c
; Sequence 1, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
;  APPLICANT: Horvitz, H. Robert
;  APPLICANT: Hengartner, Michael
;  TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
;  TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: 01997/201005
;  CURRENT APPLICATION NUMBER: US/09/234,186
;  CURRENT FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: 07/898,933
;  EARLIER FILING DATE: 1992-06-12
;  EARLIER APPLICATION NUMBER: 07/927,681
;  EARLIER FILING DATE: 1992-08-10
;  EARLIER APPLICATION NUMBER: 08/288,295
;  EARLIER FILING DATE: 1994-08-10
;  EARLIER APPLICATION NUMBER: 08/801,248
;  EARLIER FILING DATE: 1997-02-19
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 6559
;   TYPE: DNA
;   ORGANISM: Caenorrhabditis elegans
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(6559)
;   OTHER INFORMATION: n = A,T,C or G
US-09-234-186-1

  Query Match             6.5%;  Score 32.4;  DB 4;  Length 6559;
  Best Local Similarity   52.2%;  Pred. No. 4.3;
  Matches   72;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     358 GACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTC 417
           | | |  |  || || ||   |      ||   |||||  ||  | | ||  | ||||
Db    6430 GTCGACTGAAAAAGGCAGCCTAGACTGCTAGCCAACAATTTTTATACCTACAATACTTAA 6371

Qy     418 TATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTT 477
           |  | | ||||   ||||  ||| || |||| | | |    || || |   ||| |||
Db    6370 AAATGTATTTGTTTAAAGGAACACGTAGTTTTCTGAGTGGGTCTTGCCACGATCACAGTC 6311

Qy     478 TTCATTGTGTCTTAACAA 495
           |  | ||  ||  | ||
Db    6310 ATAGTGGTCGCTGCAGAA 6293


RESULT 13
US-09-233-527-1/c
; Sequence 1, Application US/09233527
; Patent No. 6465617
; GENERAL INFORMATION:
;  APPLICANT:  Horvitz, H. Robert
;  APPLICANT:  Hengartner, Michael
;  TITLE OF INVENTION:  IDENTIFICATION AND CHARACTERIZATION OF A
;  TITLE OF INVENTION:  GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  01997/201004
;  CURRENT APPLICATION NUMBER: US/09/233,527
;  CURRENT FILING DATE:  1999-01-20
;  PRIOR APPLICATION NUMBER:  07/898,933
;  PRIOR FILING DATE:  1992-06-12
;  PRIOR APPLICATION NUMBER:  07/927,681
;  PRIOR FILING DATE:  1992-08-10
;  PRIOR APPLICATION NUMBER:  08/288,295
;  PRIOR FILING DATE:  1994-08-10
;  PRIOR APPLICATION NUMBER:  08/801,248
;  PRIOR FILING DATE:  1997-02-19
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 6559
;   TYPE: DNA
;   ORGANISM: Caenorrhabditis elegans
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(6559)
;   OTHER INFORMATION: n = A,T,C or G
US-09-233-527-1

  Query Match             6.5%;  Score 32.4;  DB 4;  Length 6559;
  Best Local Similarity   52.2%;  Pred. No. 4.3;
  Matches   72;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     358 GACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTC 417
           | | |  |  || || ||   |      ||   |||||  ||  | | ||  | ||||
Db    6430 GTCGACTGAAAAAGGCAGCCTAGACTGCTAGCCAACAATTTTTATACCTACAATACTTAA 6371

Qy     418 TATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTT 477
           |  | | ||||   ||||  ||| || |||| | | |    || || |   ||| |||
Db    6370 AAATGTATTTGTTTAAAGGAACACGTAGTTTTCTGAGTGGGTCTTGCCACGATCACAGTC 6311

Qy     478 TTCATTGTGTCTTAACAA 495
           |  | ||  ||  | ||
Db    6310 ATAGTGGTCGCTGCAGAA 6293


RESULT 14
```

```
PCT-US93-05651-1/c
; Sequence 1, Application PC/TUS9305651
;  GENERAL INFORMATION:
;    TITLE OF INVENTION:  A Gene Which Prevents Programmed Cell Death
;    NUMBER OF SEQUENCES:  5
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  diskette
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/05651
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6560 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US93-05651-1

  Query Match             6.5%;  Score 32.4;  DB 5;  Length 6560;
  Best Local Similarity   52.2%;  Pred. No. 4.3;
  Matches   72;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     358 GACCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTC 417
           | | |  |  || || ||   |       ||   |||||  ||  | | ||  | ||||
Db    6431 GTCGACTGAAAAAGGCAGCCTAGACTGCTAGCCAACAATTTTTATACCTACAATACTTAA 6372

Qy     418 TATAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAAGTT 477
            |  | | ||||   ||||  ||| || |||| | | |    || || |   ||| |||
Db    6371 AAATGTATTTGTTTAAAGGAACACGTAGTTTTCTGAGTGGGTCTTGCCACGATCACAGTC 6312

Qy     478 TTCATTGTGTCTTAACAA 495
           |  | ||  ||  | ||
Db    6311 ATAGTGGTCGCTGCAGAA 6294


RESULT 15
US-08-471-033-39
; Sequence 39, Application US/08471033
; Patent No. 5770696
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5770696el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,033
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  P-40,403
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV7 - SQLv3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8582
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1241 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  9..1238
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence encoding VIP2A(a) with the Bacillus secretion sig
;      OTHER INFORMATION:  removed as contained in pCIB5527"
US-08-471-033-39

  Query Match             6.4%;  Score 32;  DB 1;  Length 1241;
  Best Local Similarity   51.4%;  Pred. No. 2.7;
  Matches   74;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy      59 ACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTT 118
           ||   || ||  |  || | || || ||  ||||||||| ||| ||  |        |
Db     139 ACTTCCTGGACAACAAGAACGACATCAAGACCAACTACAAGGAGATCACCTTCAGCATAG 198

Qy     119 CTGAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATT 178
           | |  | | |  | ||  |||||     |||| ||||||     |   | ||    ||
Db     199 CCGGCAGCTTCGAGGACGAGATCAAGGACCTGAAGGAGATCGACAAGATGTTCGACAAGA 258

Qy     179 CCCTTCACATCTTCAACGCCATCA 202
           ||   |  | |  || |  |||||
Db     259 CCAACCTGAGCAACAGCATCATCA 282


Search completed: February  7, 2003, 09:06:22
Job time : 63.6858 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 07:01:14 ; Search time 32.3549 Seconds
(without alignments)
7339.716 Million cell updates/sec

Title: US-09-687-837-1_COPY_3800_4300
Perfect score: 501
Sequence: 1 caatttgtggaagcttgcgg.....ttgtgtcttaacaaaggtgt 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	336	67.1	487	10	US-09-998-598-1612	Sequence 1612, Ap
	2	139	27.7	1031	10	US-09-876-889-157	Sequence 157, App
	3	71	14.2	423	10	US-09-960-352-3043	Sequence 3043, Ap
	4	67.4	13.5	4391	10	US-09-736-969A-7	Sequence 7, Appli
	5	67.4	13.5	6454	10	US-09-736-969A-1	Sequence 1, Appli
	6	50	10.0	434	9	US-09-796-692-4813	Sequence 4813, Ap
	7	44	8.8	505	9	US-09-796-692-5039	Sequence 5039, Ap
	8	44	8.8	505	9	US-09-796-692-5223	Sequence 5223, Ap
	9	41.8	8.3	244	10	US-09-736-960-84	Sequence 84, Appl
	10	41.8	8.3	4026	10	US-09-736-960-3	Sequence 3, Appli
	11	41.8	8.3	7215	10	US-09-736-960-1	Sequence 1, Appli
	12	41.8	8.3	66686	10	US-09-736-960-86	Sequence 86, Appl
	13	37.2	7.4	521	10	US-09-964-824A-1	Sequence 1, Appli
	14	37.2	7.4	521	10	US-09-954-456-1193	Sequence 1193, Ap
	15	36.6	7.3	696	10	US-09-815-242-7900	Sequence 7900, Ap
c	16	36.4	7.3	143068	10	US-09-967-768A-316	Sequence 316, App
	17	35.2	7.0	126512	10	US-09-804-474A-3	Sequence 3, Appli
	18	35	7.0	23626	10	US-09-764-878-261	Sequence 261, App
	19	35	7.0	23626	10	US-09-764-860-940	Sequence 940, App
	20	35	7.0	23632	10	US-09-764-878-262	Sequence 262, App
	21	35	7.0	23632	10	US-09-764-860-941	Sequence 941, App
c	22	34.8	6.9	429	10	US-09-800-729-160	Sequence 160, App
c	23	34.8	6.9	690	10	US-09-815-242-7864	Sequence 7864, Ap
c	24	34.8	6.9	368004	10	US-09-949-654-3	Sequence 3, Appli
	25	32.8	6.5	1461	10	US-09-880-192-42	Sequence 42, Appl
c	26	32.8	6.5	3891	12	US-10-051-952-1	Sequence 1, Appli
c	27	32.8	6.5	4835	10	US-09-288-326-10	Sequence 10, Appl
	28	32.6	6.5	245	10	US-09-864-761-23212	Sequence 23212, A
	29	32.6	6.5	469	10	US-09-864-761-6499	Sequence 6499, Ap
	30	32.6	6.5	563	10	US-09-833-381-542	Sequence 542, App
	31	32.6	6.5	4753	9	US-10-098-841-18	Sequence 18, Appl
	32	32.4	6.5	695	9	US-10-016-157A-69	Sequence 69, Appl
c	33	32	6.4	1767	10	US-09-864-761-30592	Sequence 30592, A
	34	32	6.4	2604	10	US-09-529-063-9	Sequence 9, Appli
	35	31.8	6.3	384	10	US-09-983-965-198	Sequence 198, App
	36	31.8	6.3	1750	10	US-09-925-300-633	Sequence 633, App
	37	31.6	6.3	337	10	US-09-864-761-22871	Sequence 22871, A
	38	31.4	6.3	740	10	US-09-735-705-6	Sequence 6, Appli
	39	31.4	6.3	740	10	US-09-850-716A-6	Sequence 6, Appli
	40	31.4	6.3	740	10	US-09-897-778-6	Sequence 6, Appli
c	41	31.4	6.3	1249	10	US-09-925-301-501	Sequence 501, App
	42	31.4	6.3	4218	10	US-09-752-639-8	Sequence 8, Appli
	43	31.4	6.3	4218	10	US-09-984-198-8	Sequence 8, Appli
c	44	31.4	6.3	202001	10	US-09-734-674-3	Sequence 3, Appli
c	45	31.2	6.2	2000	9	US-09-938-842A-4943	Sequence 4943, Ap

ALIGNMENTS

RESULT 1
US-09-998-598-1612
; Sequence 1612, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 1612
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-1612

Query Match 67.1%; Score 336; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.2e-91;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    152 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 211

Qy     61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 271

Qy    121 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    272 GAAATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCC 331

Qy    181 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    332 CTTCACATCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACC 391
```

```
Qy    241 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    392 AGCTCGTCTTCGGTCGTGTGATTACATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTCA 451

Qy    301 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACT 336
          ||||||||||||||||||||||||||||||||||
Db    452 TTTGCAAACTCAGGATGCTTTCCAAAGCCAATCACT 487


RESULT 2
US-09-876-889-157
; Sequence 157, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
;  TITLE OF INVENTION:  CANCER THERAPY AND DIAGNOSIS
;  FILE REFERENCE: 210121.466C3
;  CURRENT APPLICATION NUMBER: US/09/876,889
;  CURRENT FILING DATE:  2001-06-06
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 157
;   LENGTH: 1031
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-876-889-157

  Query Match             27.7%;  Score 139;  DB 10;  Length 1031;
  Best Local Similarity   96.6%;  Pred. No. 5.1e-32;
  Matches  142;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 570

Qy     61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 630

Qy    121 GAAATCATGCATGAGCAGATCTGCCCC 147
          |||||||||||||||| |  |  ||
Db    631 GAAATCATGCATGAGCAGGTGAGGGCC 657


RESULT 3
US-09-960-352-3043
; Sequence 3043, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3043
;   LENGTH: 423
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 14-BOVMS1-017-Q1-E1-D5
US-09-960-352-3043

  Query Match             14.2%;  Score 71;  DB 10;  Length 423;
  Best Local Similarity   70.4%;  Pred. No. 8.9e-12;
  Matches   95;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy      1 CAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGAC 60
          |||||||  || || || || || ||| | |  || || || || || || || || || ||
Db    289 CAATTTGCAGATGCGTGTGGCCAGGCCCTGGATGTGAATGAGCGCCTCATCAAGGAGGAT 348

Qy     61 CAGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCT 120
          ||||| ||||| ||||| ||| |||   || | |||| ||| |||   |  ||||| ||
Db    349 CAGCTGGAGTACCAGGAGGAACTGAGGTCCCATTACAAGGACATGCTCAGCGAGCTCTCG 408

Qy    121 GAAATCATGCATGAG 135
          | |||||||  |||||
Db    409 GGAATCATGAATGAG 423


RESULT 4
US-09-736-969A-7
; Sequence 7, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-4 Transmembrane Protein
;  FILE REFERENCE: 020054-000411US
;  CURRENT APPLICATION NUMBER: US/09/736,969A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 4391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human CLASP-4 cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (414)..(4058)
;   OTHER INFORMATION: human CLASP-4
US-09-736-969A-7

  Query Match             13.5%;  Score 67.4;  DB 10;  Length 4391;
  Best Local Similarity   67.4%;  Pred. No. 3.5e-10;
```

```
  Matches   95;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy      2 AATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACC 61
          ||||| |  |||| ||| |    ||  | |   |||| || || || ||||||||||| |
Db   3802 AATTTATACAAGCATGCAGCATTGCACTTGAACTAAATGAGCGGCTAATTAAAGAAGATC 3861

Qy     62 AGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTG 121
          |  | ||||| || ||||   | ||  | || | ||| || |||| || ||  | ||||
Db   3862 AAGTTGAGTACCATGAAGGGCTAAAGTCAAATTTCAGAGACATGGTAAAAGAATTATCTG 3921

Qy    122 AAATCATGCATGAGCAGATCT 142
          | || || ||||||||||| |
Db   3922 ACATTATCCATGAGCAGATAT 3942


RESULT 5
US-09-736-969A-1
; Sequence 1, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-4 Transmembrane Protein
;  FILE REFERENCE: 020054-000411US
;  CURRENT APPLICATION NUMBER: US/09/736,969A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 6454
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-4 cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (95)..(6121)
;   OTHER INFORMATION: human CLASP-4
US-09-736-969A-1

  Query Match             13.5%;  Score 67.4;  DB 10;  Length 6454;
  Best Local Similarity   67.4%;  Pred. No. 4.3e-10;
  Matches   95;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy      2 AATTTGTGGAAGCTTGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACC 61
          ||||| |  |||| ||| |    ||  | |   |||| || || || ||||||||||| |
Db   5865 AATTTATACAAGCATGCAGCATTGCACTTGAACTAAATGAGCGGCTAATTAAAGAAGATC 5924

Qy     62 AGCTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTG 121
          |  | ||||| || ||||   | ||  | || | ||| || |||| || ||  | ||||
Db   5925 AAGTTGAGTACCATGAAGGGCTAAAGTCAAATTTCAGAGACATGGTAAAAGAATTATCTG 5984

Qy    122 AAATCATGCATGAGCAGATCT 142
          | || || ||||||||||| |
Db   5985 ACATTATCCATGAGCAGATAT 6005


RESULT 6
US-09-796-692-4813
; Sequence 4813, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4813
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-4813

  Query Match             10.0%;  Score 50;  DB 9;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-05;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    452 AAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 501
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAATCTCACTGGCAATATTCAAAGTTTTCATTGTGTCTTAACAAAGGTGT 50


RESULT 7
US-09-796-692-5039
; Sequence 5039, Application US/09796692
```

```
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5039
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (341)
;   OTHER INFORMATION: n=A,T,C or G
US-09-796-692-5039

  Query Match             8.8%;  Score 44;  DB 9;  Length 505;
  Best Local Similarity   65.0%;  Pred. No. 0.0013;
  Matches   65;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy      64 CTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAA 123
           || ||||| ||||||||| |||   || |||||| ||| |||   |  || || ||   |
Db       1 CTGGAGTACCAGGAAGAACTGAGGTCCCACTACAAGGACATGCTCAGCGAACTCTCCACA 60

Qy     124 ATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGA 163
           ||||| ||||||||||      |  ||| ||   | | |
Db      61 GTCATGAATGAGCAGATTACGGGCAGGGACGACCTGTCAA 100


RESULT 8
US-09-796-692-5223
; Sequence 5223, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5223
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-5223

  Query Match             8.8%;  Score 44;  DB 9;  Length 505;
  Best Local Similarity   65.0%;  Pred. No. 0.0013;
  Matches   65;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy      64 CTCGAGTATCAGGAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAA 123
           || ||||| ||||||||| |||   || |||||| ||| |||   |  || || ||   |
Db       1 CTGGAGTACCAGGAAGAACTGAGGTCCCACTACAAGGACATGCTCAGCGAACTCTCCACA 60

Qy     124 ATCATGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGA 163
           ||||| ||||||||||      |  ||| ||   | | |
Db      61 GTCATGAATGAGCAGATTACGGGCAGGGACGACCTGTCAA 100


RESULT 9
US-09-736-960-84
; Sequence 84, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
```

```
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
;   LENGTH: 244
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: 21st exon
US-09-736-960-84

  Query Match             8.3%;  Score 41.8;  DB 10;  Length 244;
  Best Local Similarity   56.0%;  Pred. No. 0.004;
  Matches   79;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      16 TGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAG 75
           || ||| ||||  ||| |  |||| | ||||| || |  | |||||||   || ||||||
Db      38 TGTGGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAGACCAGAGGGAATATCAG 97

Qy      76 GAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAG 135
           | ||| | |||   ||||| |   |  |     || | ||      ||| ||  |   |
Db      98 CAGGAACTCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGCGG 157

Qy     136 CAGATCTGCCCCCTGGAGGAG 156
           | ||       ||| |  ||
Db     158 AAAATTCCAGAACTGTACAAG 178


RESULT 10
US-09-736-960-3
; Sequence 3, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 4026
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: preliminary human CLASP-5 cDNA sequence
;   NAME/KEY: CDS
;   LOCATION: (1)..(3066)
;   OTHER INFORMATION: human CLASP-5
US-09-736-960-3

  Query Match             8.3%;  Score 41.8;  DB 10;  Length 4026;
  Best Local Similarity   56.0%;  Pred. No. 0.017;
  Matches   79;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      16 TGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAG 75
           || ||| ||||  ||| |  |||| | ||||| || |  | |||||||   || ||||||
Db    2740 TGTGGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAGACCAGAGGGAATATCAG 2799

Qy      76 GAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAG 135
           | ||| | |||   ||||| |   |  |     || | ||      ||| ||  |   |
Db    2800 CAGGAACTCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGCGG 2859

Qy     136 CAGATCTGCCCCCTGGAGGAG 156
           | ||       ||| |  ||
Db    2860 AAAATTCCAGAACTGTACAAG 2880


RESULT 11
US-09-736-960-1
; Sequence 1, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
```

```
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 7215
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length human CLASP-5 cDNA
;   NAME/KEY: CDS
;   LOCATION: (112)..(6159)
;   OTHER INFORMATION: human CLASP-5
US-09-736-960-1

  Query Match             8.3%;  Score 41.8;  DB 10;  Length 7215;
  Best Local Similarity   56.0%;  Pred. No. 0.022;
  Matches   79;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      16 TGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAG 75
           || ||| ||||  ||| |  |||| | ||||| || |  | |||||||   || ||||||
Db    5929 TGTGGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAGACCAGAGGGAATATCAG 5988

Qy      76 GAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAG 135
            | ||| | |||   ||||| |   |  |     || | ||      ||| ||  |   |
Db    5989 CAGGAACTCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGCGG 6048

Qy     136 CAGATCTGCCCCCTGGAGGAG 156
           | ||       ||| |  ||
Db    6049 AAAATTCCAGAACTGTACAAG 6069


RESULT 12
US-09-736-960-86
; Sequence 86, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-5 Transmembrane Protein
;  FILE REFERENCE: 020054-000511US
;  CURRENT APPLICATION NUMBER: US/09/736,960
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
;   LENGTH: 66686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: ordered human genomic DNA at CLASP-5 locus
US-09-736-960-86

  Query Match             8.3%;  Score 41.8;  DB 10;  Length 66686;
  Best Local Similarity   56.0%;  Pred. No. 0.068;
  Matches   79;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      16 TGCGGTCAAGCCTTAGCGGTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAG 75
           || ||| ||||  ||| |  |||| | ||||| || |  | |||||||   || ||||||
Db   62399 TGTGGTGAAGCTGTAGAGAAAAACAAGCGTCTCATCACGGCAGACCAGAGGGAATATCAG 62458

Qy      76 GAAGAAATGAAAGCCAACTACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAG 135
            | ||| | |||   ||||| |   |  |     || | ||      ||| ||  |   |
Db   62459 CAGGAACTCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGCGG 62518

Qy     136 CAGATCTGCCCCCTGGAGGAG 156
           | ||       ||| |  ||
Db   62519 AAAATTCCAGAACTGTACAAG 62539


RESULT 13
US-09-964-824A-1
; Sequence 1, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
;  APPLICANT: Horrigan, Stephen
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-73
;  CURRENT APPLICATION NUMBER: US/09/964,824A
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: US/60/236,033
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,032
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,028
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 583
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 521
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(521)
;   OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-1

  Query Match             7.4%;  Score 37.2;  DB 10;  Length 521;
  Best Local Similarity   57.4%;  Pred. No. 0.14;
  Matches   66;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     360 CCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTA 419
           ||  ||||  |||||| | |  | | |||| || ||  |||       | ||  |   ||
```

```
Db     245 CCTGGGGGTTGGGGAGGGGAGTGGATTAAATAAAAAGTTTAGAAGGCCATAGNATAAATA 304

Qy     420 TAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAA 474
           | | | | ||| |||  |  | ||||  |||||||       ||||||  | |||
Db     305 TCGAAATAGTATGAATTTTAATATATACTTTTAAAGGGGTTAGGCAATGATGAAA 359


RESULT 14
US-09-954-456-1193
; Sequence 1193, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1193
;   LENGTH: 521
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1193

  Query Match             7.4%;  Score 37.2;  DB 10;  Length 521;
  Best Local Similarity   57.4%;  Pred. No. 0.14;
  Matches   66;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     360 CCAAGGGGAAGGGGAGAGAAAGGAAATAAAGAACAACGTTATTTCTTAACAGACTTTCTA 419
           ||  ||||  ||||||  | |  | | |||| || ||  |||       | ||  |   ||
Db     245 CCTGGGGGTTGGGGAGGGGAGTGGATTAAATAAAAAGTTTAGAAGGCCATAGNATAAATA 304

Qy     420 TAGGAGTTGTAAGAAGGTGCACATATTTTTTTAAATCTCACTGGCAATATTCAAA 474
           | | | | ||| |||  |  | ||||  |||||||       ||||||  | |||
Db     305 TCGAAATAGTATGAATTTTAATATATACTTTTAAAGGGGTTAGGCAATGATGAAA 359


RESULT 15
US-09-815-242-7900
; Sequence 7900, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7900
;   LENGTH: 696
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(696)
US-09-815-242-7900

  Query Match             7.3%;  Score 36.6;  DB 10;  Length 696;
  Best Local Similarity   52.3%;  Pred. No. 0.25;
  Matches   81;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy     128 TGCATGAGCAGATCTGCCCCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACA 187
           |||   ||| | |||| |  | || |  ||  | | | || |   |||  |  || |  |
Db     249 TGCCGAAGCCGCTCTGGCTGCCGGCGCCGACAAGGTGGGTATGGACGAGCTGGCTGCCGA 308

Qy     188 TCTTCAACGCCATCAGTGGGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGT 247
             |  || ||    |   | ||| | ||     ||  | ||| | ||||| |  ||  ||
Db     309 AATGAAAGGCGGCGACCTGAACTACGACGTCGTCATCGCTTCCCCGGATGCCATGCGTGT 368

Qy     248 CTTCGGTCGTGTGATTACATCTCATGGCCCGTGTG 282
           | ||||||   ||          |  |||||| |||
Db     369 CGTCGGTCAGCTGGGCCAGATCCTCGGCCCGCGTG 403


Search completed: February  7, 2003, 09:15:32
Job time : 83.3549 secs
```